



Razonamiento estadístico

Bennett Briggs Triola



Razonamiento estadístico

Jeffrey O. Bennett
University of Colorado at Boulder

William L. Briggs
University of Colorado at Denver

Mario F. Triola
Dutchess Community College

TRADUCCIÓN
Victor Hugo Ibarra Mercado
Escuela de Actuaría
Universidad Anahuac, México

REVISIÓN TÉCNICA
Roberto Hernández Ramírez
Departamento de Matemáticas
Universidad de Monterrey, Nuevo León

PEARSON

Bennett, Jeffrey O., Briggs, William L., Triola, Mario F.

Razonamiento estadístico

Primera edición

PEARSON EDUCACIÓN, México, 2011

ISBN: 978-607-32-0759-1

Área: Matemáticas

Formato: 21 × 27 cm

Páginas: 512

Authorized translation from the English language edition, entitled: *Statistical reasoning for everyday life*, 3rd Edition, by Jeffrey O. Bennett, William L. Briggs, Mario F. Triola, published by Pearson Education Inc., publishing as Pearson, Copyright © 2009. Original ISBN 978-0-321-28672-7. All rights reserved.

Traducción autorizada de la edición en idioma inglés, *Statistical reasoning for everyday life* por Jeffrey O. Bennett, William L. Briggs, Mario F. Triola, publicada por Pearson Education Inc., publicada como Pearson, Copyright © 2011. Todos los derechos reservados.

Esta edición en español es la única autorizada.

Edición en español

Editor: Carlos Mario Ramírez Torres
carlosmario.ramirez@pearson.com
Editor de desarrollo: Claudia Silva Morales
Supervisor de producción: Juan José García Guzmán

Edición en inglés

Editor in Chief: Deirdre Lynch
Associate Editor: Sara Oliver Gordus
Editorial Assistant: Christina Lepre
Senior Managing Editor: Karen Wernholm
Senior Production Supervisor: Tracy Patruno
Senior Designer: Barbara T. Atkinson
Photo Researcher: Beth Anderson
Digital Assets Manager: Marianne Groth
Senior Media Producer: Cecilia Fleming
Software Development: Edward Chappell and Mary Wright
Marketing Manager: Wayne Parkins
Marketing Assistants: Caroline Celano and Kathleen DeChavez
Senior Author Support/Technology Specialist: Joe Vetere
Senior Prepress Supervisor: Caroline Fell
Rights and Permissions Advisor: Dana Weightman
Manufacturing Manager: Evelyn Beaton
Text Design: Leslie Haimes
Production Coordination: Lifland et al., Bookmakers
Composition: Progressive Information Technologies
Illustrations: Scientific Illustrators
Cover photos: Shutterstock

PRIMERA EDICIÓN, 2011.

D.R.© 2011 por Pearson Educación de México, S.A. de C.V.
Atacomulco 500 – 5o. piso
Industrial Atoto, 53519 Naucalpan de Juárez, Estado de México

Cámara Nacional de la Industria Editorial Mexicana. Reg. núm. 1031.

Custom Publishing es una marca registrada de Pearson Educación de México, S.A. de C.V.

Reservados todos los derechos. Ni la totalidad ni parte de esta publicación pueden reproducirse, registrarse o transmitirse, por un sistema de recuperación de información, en ninguna forma ni por ningún medio, sea electrónico, mecánico, fotoquímico, magnético o electroóptico, por fotocopia, grabación o cualquier otro, sin permiso previo por escrito del editor.

El préstamo, alquiler o cualquier otra forma de cesión de uso de este ejemplar requerirá también la autorización del editor o de sus representantes.

ISBN VERSIÓN IMPRESA: 978-607-32-0759-1
ISBN VERSIÓN E-BOOK: 978-607-32-0760-7
ISBN VERSIÓN E-CHAPTER: 978-607-32-0761-4

Impreso en México. *Printed in Mexico*

1 2 3 4 5 6 7 8 9 10 –14 13 12 11

Este libro está dedicado a todos aquellos
quienes tratan de hacer de este mundo
un lugar mejor. Esperamos que el
estudio de la estadística te sea de utilidad
en todos tus esfuerzos por lograrlo.

También está dedicado a quienes
hacen nuestras vidas más brillantes,
especialmente a Lisa, Grant, Brooke, Julie,
Katie, Ginny, Marc y Scott.

Contenido

Prefacio vii

Complementos xi

Al estudiante: Cómo tener éxito en su curso de estadística xiii

Índice de aplicaciones xv

Capítulo 1 Hablemos de estadística 1

1.1 ¿Qué es la estadística? 2

1.2 Muestreo 11

1.3 Tipos de estudios estadísticos 21

1.4 ¿Debe tener confianza en una investigación estadística? 34

HABLEMOS DE SOCIOLOGÍA: ¿Las guarderías generan niños agresivos? 48

HABLEMOS DE SALUD PÚBLICA: ¿Es saludable su estilo de vida? 51

Capítulo 2 Medición en estadística 53

2.1 Tipos de datos y niveles de medida 54

2.2 Manejo de errores 59

2.3 Uso de porcentajes en estadística 67

2.4 Números índice 75

HABLEMOS DE POLÍTICA: ¿Quién se beneficia con una reducción de impuestos? 84

HABLEMOS DE ECONOMÍA: ¿Está mejorando nuestro nivel de vida? 86

Capítulo 3 Exhibición visual de datos 89

3.1 Tablas de frecuencia 90

3.2 Distribución gráfica de datos 99

3.3 Gráficas en los medios 112

3.4 Algunas precauciones con respecto a gráficas 125

USO DE LA TECNOLOGÍA 136

HABLEMOS DE HISTORIA: ¿La guerra puede describirse con una gráfica? 137

HABLEMOS DE MEDIO AMBIENTE: ¿Cómo podemos visualizar el calentamiento global? 140

Capítulo 4 Descripción de datos 145

4.1 ¿Qué es un promedio? 146

4.2 Formas de las distribuciones 157

4.3 Medidas de variación 164

4.4 Paradojas estadísticas 178

USO DE LA TECNOLOGÍA 187

HABLEMOS DE LA BOLSA DE VALORES: ¿Qué es promedio cuando se habla del Dow Jones? 188

HABLEMOS DE ECONOMÍA: ¿Los ricos se vuelven más ricos? 191

Capítulo 5	Un mundo normal	195
5.1	¿Qué es normal?	196
5.2	Propiedades de la distribución normal	204
5.3	El teorema del límite central	215
	USO DE LA TECNOLOGÍA	225
	HABLEMOS DE EDUCACIÓN: ¿Qué podemos aprender de las tendencias del SAT?	226
	HABLEMOS DE PSICOLOGÍA: ¿Somos más inteligentes que nuestros padres?	229
Capítulo 6	Probabilidad en estadística	233
6.1	El papel de la probabilidad en estadística: significancia estadística	234
6.2	Fundamentos de probabilidad	238
6.3	Probabilidad con grandes números	251
6.4	Ideas de riesgo y esperanza de vida	260
6.5	Combinación de probabilidades (sección complementaria)	267
	HABLEMOS DE CIENCIAS SOCIALES: ¿Son justas las loterías?	279
	HABLEMOS DE LEYES: ¿El ADN es una huella confiable?	282
Capítulo 7	Correlación y causalidad	285
7.1	Búsqueda de correlación	286
7.2	Interpretación de correlaciones	299
7.3	Rectas de mejor ajuste y pronóstico	307
7.4	La búsqueda de la causalidad	315
	USO DE LA TECNOLOGÍA	324
	HABLEMOS DE EDUCACIÓN: ¿Qué ayuda a los niños a aprender a leer?	325
	HABLEMOS DE MEDIO AMBIENTE: ¿Qué causa el calentamiento global?	328
Capítulo 8	De muestras a poblaciones	333
8.1	Distribuciones de muestreo	334
8.2	Estimación de medias poblacionales	346
8.3	Estimación de proporciones poblacionales	355
	USO DE LA TECNOLOGÍA	362
	HABLEMOS DE HISTORIA: ¿Dónde inició la estadística?	363
	HABLEMOS DE LITERATURA: ¿Cuántas palabras conoció Shakespeare?	366
Capítulo 9	Pruebas de hipótesis	369
9.1	Fundamentos de las pruebas de hipótesis	370
9.2	Pruebas de hipótesis para medias poblacionales	380
9.3	Pruebas de hipótesis para proporciones poblacionales	394
	USO DE LA TECNOLOGÍA	402
	HABLEMOS DE SALUD Y EDUCACIÓN: ¿Su educación le ayudará a vivir más tiempo?	403
	HABLEMOS DE AGRICULTURA: ¿Los alimentos modificados genéticamente son seguros?	406

Capítulo 10	Pruebas t, tablas de dos entradas y ANOVA	409
10.1	La distribución t para inferencias acerca de una media	410
10.2	Pruebas de hipótesis con tablas de dos colas	417
10.3	Análisis de varianza (ANOVA de un factor)	429
	USO DE LA TECNOLOGÍA	437
	HABLEMOS DE CRIMINOLOGÍA: ¿Puede descubrir un fraude cuando lo ve?	439
	HABLEMOS DE EDUCACIÓN: ¿Qué puede hacer una alumna de cuarto grado con estadística?	443
Epílogo:	una perspectiva sobre estadística	445
Apéndice A:	tablas de puntuación z	446
Apéndice B:	tabla de números aleatorios	449
Lecturas sugeridas		450
Créditos		452
Glosario		454
Respuestas	R-1	
Índice	I-1	

Prefacio

¿Por qué estudiar estadística?

El escritor de ciencia ficción H. G. Wells, una vez escribió, “El razonamiento estadístico algún día será tan necesario para el ciudadano eficiente como la capacidad para leer y escribir”. El futuro que Wells imaginó está aquí. Ahora, la estadística es una parte importante de la vida diaria, inevitablemente si usted inicia un nuevo negocio, si decide cómo planear su futuro financiero, o sólo observar las noticias en la televisión. La estadística surge en todo, desde encuestas de opinión a reportes económicos y hasta la última investigación sobre la prevención de cáncer. Por tanto, la comprensión de las ideas centrales en que se basa la estadística es crucial para su éxito en el mundo moderno.

¿Qué clase de estadística aprenderá en este libro?

La estadística es un campo rico de estudio —tan rico que es posible estudiarla durante toda una vida y aún sentir que queda mucho por aprender—. Sin embargo, puede entender las ideas centrales de estadística con sólo un trimestre o un semestre de estudio académico. Este libro está diseñado para ayudarle a aprender estas ideas centrales. Las ideas que estudiará en este libro representan la estadística que *necesitará* en su vida diaria —y que de manera razonable puede aprender en un curso de estudio—. En particular, hemos diseñado este libro con tres propósitos específicos:

1. Para proporcionarle la comprensión de la estadística que necesitará en cursos **universitarios**, en particular en ciencias sociales tal como economía, psicología, sociología y ciencias políticas.
3. Para ayudarle a desarrollar la capacidad de razonar usando información estadística —una capacidad que es crucial para casi cualquier **carrera** en el mundo moderno.
3. Para proporcionarle el poder de evaluar la gran cantidad de informes noticiosos de estudios estadísticos que encontrará en su **vida** diaria, de este modo le ayudará a formar opiniones acerca de sus conclusiones y para decidir si las conclusiones deben influir la forma de su vida.

¿Quién debe leer este libro?

Esperamos que este libro será útil para todos, pero está diseñado principalmente para estudiantes que *no* planean continuar con cursos avanzados de estadística. En particular, este libro debe proporcionar una introducción adecuada a estadística para estudiantes que se especializan en un amplio rango de campos que requieren dominio estadístico, incluyendo la mayoría de las humanidades y ciencias sociales. El nivel de este texto debe ser adecuado para cualquiera que haya completado dos años de matemáticas de preparatoria.

Enfoque

Este libro toma un enfoque diseñado para ayudarle a entender las ideas estadísticas importantes de manera cualitativa, usando técnicas cuantitativas sólo cuando clarifiquen esas ideas. A continuación se proporcionan unas cuantas de las estrategias pedagógicas clave que guiaron la creación de este libro.

INICIAR CON EL PANORAMA GENERAL. La mayoría de las personas ingresan a un curso de estadística teniendo pocos conocimientos previos del tema, de modo que es importante mantener la vista del propósito global de la estadística mientras se aprenden las ideas o métodos individuales. Por tanto, iniciamos este libro con un panorama amplio de la estadística en el capítulo 1, en el que explicamos la relación entre muestras y poblaciones, analizamos los métodos

de muestreo y los diferentes tipos de estudios estadísticos, y mostramos numerosos ejemplos diseñados para ayudarle a decidir si crear un estudio estadístico. Este “panorama general” de la estadística proporciona un sólido fundamento para el estudio a mayor profundidad de las ideas estadísticas del resto del libro.

CONSTRUIR IDEAS PASO A PASO. El objetivo de este curso en estadística es ayudar a los estudiantes a entender los temas reales de estadística. Sin embargo, con frecuencia es más fácil iniciar mediante la investigación de ejemplos sencillos para construir el conocimiento paso a paso que entonces pueda aplicarse a estudios más complejos. Aplicamos esta estrategia en cada sección, construyendo en forma gradual hacia ejemplos reales y casos de estudio.

USO DE CÁLCULOS PARA AUMENTAR LA COMPRESIÓN. El objetivo principal de este libro es ayudar a los estudiantes a comprender las ideas y los conceptos estadísticos, pero creemos firmemente que este objetivo se alcanza mejor haciendo al menos algunos cálculos. Por tanto, incluimos técnicas computacionales siempre que ellas aumenten la comprensión de las ideas subyacentes.

RELACIÓN ENTRE PROBABILIDAD Y ESTADÍSTICA. Muchos cursos de estadística incluyen información de probabilidad, pero a los estudiantes el concepto de probabilidad con frecuencia les parece desconectado del resto de los temas. Esto es una pena, ya que la probabilidad desempeña un papel integral en la ciencia estadística. Analizamos este punto al inicio del capítulo 1, con la estructura básica de estudios estadísticos, y luego los revisamos a lo largo del libro —en particular en el capítulo 6, donde presentamos muchas ideas de probabilidad. Para aquellos cursos en los que la cobertura de probabilidad no se resalta, el capítulo 6 está diseñado para ser opcional.

MANTENER EL OBJETIVO: APLICACIÓN DEL RAZONAMIENTO ESTADÍSTICO PARA LA VIDA DIARIA. Puesto que la estadística es un tema tan rico, puede ser difícil decidir qué tan profundo ir con cualquier tema estadístico en particular. Para tomar tales decisiones para este libro, siempre regresamos al objetivo reflejado en el título. Este libro se supone que le ayuda con el razonamiento estadístico necesario en la vida diaria. Si consideramos que un tema no se encontraba con frecuencia en la vida diaria, lo dejamos fuera. En el mismo espíritu, incluimos algunos temas —tal como un análisis de porcentajes en el capítulo 2 y un estudio a profundidad de gráficas en el capítulo 3— que no se tratan con frecuencia en cursos de estadística pero son una parte importante de la estadística encontrada en la vida diaria.

Estructura modular

Aunque hemos escrito este libro de modo que pueda leerse como una narración del principio al final, reconocemos que muchos instructores podrían desear enseñar el material en un orden diferente al que hemos elegido o cubrir sólo partes seleccionadas del texto conforme el tiempo permita para clases de diferentes duraciones o con estudiantes a niveles diferentes. Por tanto, hemos organizado el libro con una estructura modular que permita a los instructores crear un curso personalizado. Los diez capítulos están organizados en general por áreas conceptuales. A su vez, cada capítulo está dividido en un conjunto de secciones autocontenidas, cada una dedicada a un tema o aplicación particular. En la mayoría de los casos, usted puede cubrir las secciones o capítulos en cualquier orden o saltar secciones que no se ajusten bien a su curso. Por favor, observe la siguiente estructura específica dentro de cada capítulo:

OBJETIVOS DE APRENDIZAJE. Cada capítulo inicia con una página que da una visión general del tema, incluyendo una lista del contenido de cada sección para los objetivos de aprendizaje.

SECCIONES NUMERADAS. Cada capítulo está subdividido en un conjunto de secciones numeradas (por ejemplo, secciones 1.1, 1.2, . . .). Para facilitar el uso de estas secciones en cualquier orden, cada sección termina con su propio conjunto de ejercicios específicos sólo para esa sección. Los ejercicios se dividen en subgrupos con encabezados que deben ser autoexplicativos, incluyendo “Alfabetización estadística y pensamiento crítico”, “Conceptos y aplicaciones”, “Proyectos para internet y más allá” y “En las noticias”. Las respuestas para casi todos los ejercicios con número impar aparecen al final del libro. Las secciones numeradas también incluyen las características pedagógicas siguientes:

- **Ejemplos y casos de estudio.** Ejemplos numerados en cada sección están diseñados para construir la comprensión y ofrecer práctica con los tipos de preguntas que aparecen en los ejercicios. Casos de estudio, que siempre se enfocan a temas reales, llevan más profundidad que los ejemplos numerados.
- **Momento de reflexión.** Las características de “Momento de reflexión” plantean preguntas conceptuales breves diseñadas para ayudar a los estudiantes a reflexionar sobre nuevas ideas importantes. También sirven como excelentes puntos de inicio para discusión en clase, en algunos casos pueden usarse para preguntas para dar mayor comprensión.

EJERCICIOS DE REPASO DEL CAPÍTULO. La parte principal de cada capítulo termina con un breve conjunto de ejercicios diseñados para ligar muchas de las ideas del capítulo. Estos ejercicios están diseñados principalmente para autoestudio, con respuestas a todas ellas en la parte final del libro.

SECCIONES HABLEMOS DE UN TEMA. Cada capítulo termina con dos secciones tituladas “Hablemos de...” que profundizan sobre temas estadísticos importantes de nuestro tiempo. Los temas de estas secciones fueron elegidos para demostrar la gran variedad de campos en los que la estadística desempeña un papel, incluyendo historia, estudios ambientales, agricultura y economía. Cada una de estas secciones “Hablemos de...” incluye un conjunto de preguntas para asignación o discusión.

Acerca de esta edición

Hemos desarrollado esta edición de *Razonamiento estadístico* con la ayuda de muchos usuarios y revisores. Además, se editó y rediseñó todo el libro para hacerlo aún más amigable para el estudiante. Para esta edición hemos realizado los cambios principales siguientes:

- Puesto que este libro está planeado para mostrar la relevancia de la estadística en la vida diaria, es crítico que las discusiones y ejemplos se hayan actualizado. Por tanto, hemos revisado o reemplazado muchos de los ejemplos del texto, los ejemplos numerados, y los casos de estudio para asegurar que reflejen los últimos datos o temas de interés. También hemos reemplazado cuatro de las veinte secciones “Hablemos de...” y actualizado todas las demás.
- Hemos vuelto a trabajar todos los conjuntos de ejercicios, reemplazado por completo 63% de los ejercicios y revisado o actualizado la información en la mayoría de los demás. Existen 1,412 ejercicios, divididos en las categorías siguientes: Aprendiendo a escribir en estadística y pensamiento crítico, Conceptos y aplicaciones, Proyectos para internet y más allá, En las noticias, Ejercicios de repaso, y —nuevo en esta edición— Cuestionario del capítulo.
- En aquellos capítulos que analizamos cálculos, la sección final numerada es seguida por una sección llamada “Uso de la tecnología”, que proporciona un breve panorama de cómo pueden hacerse los cálculos con paquetes populares como SPSS®, Excel® y STATDISK®.
- La tercera edición contiene más de treinta figuras anotadas. Las anotaciones resaltan información relevante acerca de exhibición visual de datos.
- Casi hemos reescrito por completo el capítulo 9 para proporcionar una introducción más sencilla y más enfocada a pruebas de hipótesis.
- Hemos agregado dos secciones nuevas, ambas en el capítulo 10, que cubren las distribuciones t y ANOVA de un factor. Con estas adiciones, el capítulo 10 ahora construye sobre las ideas de pruebas de hipótesis introducidas en el capítulo 9.
- Hemos movido y revisado dos secciones que aparecieron antes en el capítulo 10. La sección sobre Paradojas estadísticas (antes sección 10.2) ahora aparece al final del capítulo 4 como sección 4.4; la sección sobre Riesgo y esperanza de vida (antes sección 10.1) ahora aparece como sección 6.4.

Reconocimiento

Escribir un libro de texto requiere del esfuerzo de muchas personas además de los autores. Este libro no habría sido posible sin la ayuda de mucha gente. Queremos agradecer en particular a nuestros editores en Addison-Wesley, Greg Tobin y Deirdre Lynch, cuya confianza nos permitió crear este libro. También agradecemos al resto del equipo en Addison-Wesley quienes nos ayudaron a producir este libro, incluyendo a Sara Oliver Gordus, Christina Lepre, Tracy Patruno, Ceci Flemming, Caroline Celano, Beth Anderson, Barbara Atkinson y nuestro gerente de producción, Sally Lifland.

Para ayudar a asegurar la precisión de este texto, le damos las gracias a Matthew Bognar, Universidad de Iowa y Sheila O'Leary Weaver, Universidad de Vermont. Por la revisión de este texto y proporcionar una invaluable asesoría, agradecemos a las siguientes personas:

Revisores de esta edición

Jennifer Beineke
Western New England College

Matthew Bognar
University of Iowa

Pat Buchanan
Pennsylvania State University

Antonius H. N. Cillessen
University of Connecticut

Robert Dobrow
Carleton College

Beverly J. Ferrucci
Keene State College

Jack R. Fraenkel
San Francisco State University

Susan Janssen
University of Minnesota–Duluth

Becky Ladd
Arizona State University

Christopher Leary
SUNY Geneseo

Carrie M. Margolin
The Evergreen State College

Craig McCarthy
Ohio University

Thomas Petee
Auburn University

William S. Rayens
University of Kentucky

Pali Sen
University of North Florida

Donald Hugh Smith
Old Dominion University

Elizabeth Walters
Loyola College of Maryland

Sheila O'Leary Weaver
University of Vermont

Revisores de ediciones anteriores

Dale Bowman
University of Mississippi

Patricia Buchanan
Penn State University

Robert Buck
Western Michigan University

Olga Cordero-Brana
Arizona State University

Terry Dalton
University of Denver

Jim Daly
California Polytechnic State University

Mickle Duggan
East Central University

Juan Estrada
*Metropolitan State University,
Minneapolis–St. Paul*

Jack R. Fraenkel
San Francisco State University

Frank Grosshans
West Chester University

Silas Halperin
Syracuse University

Golde Holtzman
*Virginia Polytechnic Institute
and State University*

Colleen Kelly
San Diego State University

Jim Koehler
Stephen Lee
University of Idaho

Kung-Jong Lui
San Diego State University

Judy Marwick
Prairie State College

Richard McGrath
Penn State University

Abdelelah Mostafa
University of South Florida

Todd Ogden
University of South Carolina

Nancy Pfenning
University of Pittsburgh

Steve Rein
California Polytechnic State University

Lawrence D. Ries
University of Missouri–Columbia

Larry Ringer
Texas A&M University

John Spurrier
University of South Carolina

Gwen Terwilliger
University of Toledo

David Wallace
Ohio University

Larry Wasserman
Carnegie Mellon University

Sheila Weaver
University of Vermont

Robert Wolf
University of San Francisco

Fancher Wolfe
*Metropolitan State University,
Minneapolis–St. Paul*

Ke Wu
University of Mississippi

Complementos

COMPLEMENTOS PARA EL ESTUDIANTE

MANUAL DE SOLUCIONES DEL ESTUDIANTE. Este manual proporciona soluciones detalladas y totalmente desarrolladas para todos los ejercicios con número impar del texto y problemas de cuestionario de cada capítulo. ISBN-13: 978-0-321-28706-9; ISBN-10: 0-321-28706-1.

SITIO WEB ACOMPAÑANTE. El sitio web para este texto contiene recursos adicionales para los estudiantes, incluyendo conjuntos de datos y enlaces de la web del texto. El URL es www.aw-bc.com/bbt.

COMPLEMENTOS PARA EL INSTRUCTOR

EDICIÓN DEL INSTRUCTOR. Esta versión del texto incluye las respuestas para todos los ejercicios y problemas de los cuestionarios. (La edición del estudiante sólo contiene respuestas para los problemas con número impar).

MANUAL DE SOLUCIONES DEL INSTRUCTOR. Este manual detallado contiene soluciones para todos los ejercicios y cuestionarios de cada capítulo.

TESTGEN®. TestGen permite a los instructores construir, editar, imprimir y administrar exámenes usando un banco de preguntas computarizado desarrollado para cubrir todos los objetivos del texto. TestGen tiene bases algorítmicas, que permiten a los instructores crear versiones múltiples, pero equivalentes de la misma pregunta o examen con un clic de un botón. También los instructores pueden modificar las preguntas del banco de exámenes o agregar nuevas preguntas. Los exámenes pueden imprimirse o administrarse en línea. El software y el banco de exámenes están disponibles bajándolos del catálogo en línea de Pearson Education.

PREGUNTAS PARA APRENDIZAJE ACTIVO. Formateadas como diapositivas de PowerPoint®, estas preguntas pueden usarse con sistemas de respuesta en el salón de clases. Varias preguntas de opción múltiple están disponibles para cada sección del libro, permitiendo a los instructores evaluar con rapidez el dominio del material en el grupo. Las diapositivas están disponibles para bajarse de MyStatLab y del centro de recursos para el instructor de Pearson (www.aw-bc.com/irc).

DIAPOSITIVAS DE CLASES EN POWERPOINT®. Estas diapositivas presentan conceptos clave y definiciones del texto. Las diapositivas están disponibles para bajarse desde MyStatLab y del centro de recursos para el instructor de Pearson (www.aw-bc.com/irc).

RECURSOS DE TECNOLOGÍA

MYSTATLAB®. MyStatLab (parte de la familia de productos MyMathLab® y MathXL®) es un curso en línea específico del texto y fácil de personalizar que integra instrucción interactiva con múltiples medios con el contenido del libro de texto. Potenciado con Course-Compass™ (ambiente de enseñanza y aprendizaje en línea de Pearson Education) y con MathXL (nuestro sistema de evaluación de tareas y tutorial en línea), MyStatLab le da las herramientas que necesita para liberar todo o parte de su curso en línea, si sus estudiantes están en un laboratorio configurado o trabajan en casa. MyStatLab proporciona un conjunto de materiales para el curso que son ricos y flexibles, caracterizado por ejercicios tutoriales de respuesta abierta para práctica y dominio ilimitados. También los estudiantes pueden usar las herramientas en línea, tal como animaciones y un libro de texto en multimedios, para mejorar independientemente de su comprensión y desempeño. Los instructores pueden usar las tareas y administradores de exámenes de MyStatLab para seleccionar y asignar tareas en línea correlacionados directamente con el libro de texto, y también pueden crear y asignar sus propios ejercicios en línea e importar exámenes de TestGen para agregar flexibilidad. El registro de calificaciones en línea de MyStatLab —diseñada específicamente para matemáticas y estadística— de manera automática da un seguimiento de los resultados de las tareas y exámenes de los estudiantes y proporciona al instructor un control sobre cómo calcular la calificación final. Los instructores también pueden agregar fuera de línea (en lápiz y papel) calificaciones para el registro de calificaciones. MyStatLab también incluye acceso al Centro de Asesoría de Pearson, que proporciona a estudiantes de asesoría vía telefónica y con llamadas gratis, fax, email y sesiones interactivas en la web. MyStatLab está disponible para quienes adopten el libro. Para más información, visite nuestro sitio web en www.mystatlab.com, o contacte a su representante de ventas.

MATHXL® PARA ESTADÍSTICA. MathXL para Estadística es un sistema para tareas, tutorial y evaluación en línea que acompaña a los libros de texto de estadística de Pearson Education. Con MathXL para Estadística, los instructores pueden crear, editar y asignar tareas en línea y exámenes usando ejercicios generados de manera algorítmica correlacionados con el nivel del objetivo para el libro de texto. También pueden crear y asignar sus propios ejercicios en línea e importar exámenes de TestGen para agregar flexibilidad. Todo el trabajo de los estudiantes es seguido en el registro de calificaciones en línea de MathXL. Los estudiantes pueden presentar exámenes de capítulo en MathXL y recibir planes de estudio personalizados con base en sus resultados. El plan de estudio diagnostica las debilidades y vincula a los estudiantes directamente a ejercicios tutoriales para los

objetivos que ellos necesitan estudiar y volver a examinar. Los estudiantes también pueden tener acceso a animaciones adicionales directamente de los ejercicios seleccionados. MathXL

para Estadística está disponible para quienes adopten el libro. Para más información, visite nuestro sitio web en www.mathxl.com, o contacte a su representante de ventas.

Al estudiante

Cómo tener éxito en su curso de estadística

Si está leyendo este libro, quizás esté inscrito en un curso de estadística de algún tipo. Las claves para tener éxito en su curso incluyen abordar el material dentro de un marco mental abierto y optimista, poniendo mucha atención en la utilidad y empleo que la estadística puede tener en su vida, y hacer un estudio eficiente y eficaz. Las secciones siguientes ofrecen unas sugerencias específicas que puede utilizar cuando estudie.

CÓMO USAR ESTE LIBRO

- Antes de dar estrategias más generales para estudiar, a continuación están algunas directrices que le ayudarán a usar *este* libro de manera más eficaz.

Antes de hacer cualquiera de los ejercicios asignados, lea el material asignado *dos veces*.

- En la primera lectura, lea rápidamente para obtener “una idea” del material y los conceptos que se presentan.
- En la segunda, lea el material con mayor profundidad y trabaje completamente los ejemplos con mucho cuidado.
- Durante la segunda lectura, tome notas que le ayudarán cuando regrese a estudiar posteriormente. En particular:
 - ¡Utilice los márgenes!* Los amplios márgenes en este libro están diseñados para darle espacio suficiente para hacer notas cuando estudie.
 - No resalte —*¡subraye!* Utilice una pluma o un lápiz para subrayar el material que requiera mayor cuidado y por tanto le ayudará a mantenerse alerta cuando estudie.
- Aprenderá mejor *haciendo*, así que después de que complete la lectura asegúrese de hacer una cantidad suficiente de ejercicios del final de la sección y los ejercicios de repaso del capítulo. En particular, trate algunos de los ejercicios que tengan respuesta al final del libro, además de todos los ejercicios asignados por su instructor.
- Si tiene acceso a MyStatLab con este libro, asegúrese de aprovechar las ventajas de la gran cantidad de recursos disponibles en este sitio web.

ADMINISTRE SU TIEMPO

Una regla empírica general para clases en la universidad es que debe esperar estudiar alrededor de 2 a 3 horas por semana *fuera* de clase por cada unidad de crédito. Con base en esta regla empírica, un estudiante que toma 15 horas crédito debe esperar destinar entre 30 y 45 horas a la semana para estudiar fuera de clase. Combinadas con el tiempo de clase, este trabajo da un total de 45 a 60 horas destinadas al trabajo académico, no mucho más del tiempo requerido en un trabajo común, y usted elige su propio horario. Por supuesto, si está trabajando mientras asiste a la escuela, necesitará administrar de manera cuidadosa su tiempo. A continuación están algunas directrices de cómo podría dividir su tiempo de estudio.

Si su curso es de	Tiempo para leer el texto asignado (por semana)	Tiempo para tareas (por semana)	Tiempo para revisar y preparación de exámenes (promedio por semana)	Tiempo total de estudio (por semana)
3 créditos	1 a 2 horas	3 a 5 horas	2 horas	6 a 9 horas
4 créditos	2 a 3 horas	3 a 6 horas	3 horas	8 a 12 horas
5 créditos	2 a 4 horas	4 a 7 horas	4 horas	10 a 15 horas

Si determina que usted está destinando menos horas que las que sugieren estas directrices, quizá podría mejorar sus calificaciones estudiando más. Si está destinando más horas de las que sugieren estas directrices, podría ser que esté estudiando de manera poco eficiente; en ese caso, debe hablar con su instructor acerca de cómo estudiar de manera más eficiente.

ESTRATEGIAS GENERALES PARA ESTUDIAR

- No falte a clase. Escuchar las clases y participar en las discusiones es mucho mejor que leer los apuntes de alguien más. La participación activa le ayudará a retener lo que está aprendiendo.
- Administre cuidadosamente su tiempo. Destinar una o dos horas diarias es más efectivo, y mucho menos penoso, que estudiar toda la noche antes de hacer la tarea o antes de exámenes.
- Si un concepto le da problemas, haga una lectura adicional o resuelva más problemas que los que le han sido asignados. Y si aún tiene problemas, pida ayuda, seguramente puede encontrar amigos, compañeros o maestros que con gusto le ayudarán a que usted aprenda.
- El trabajo con amigos puede ser muy valioso y le ayuda a resolver problemas difíciles. Sin embargo, asegúrese que usted está aprendiendo *con* sus amigos y no que está dependiendo de ellos.

PREPARACIÓN PARA EXÁMENES

- Vuelva a trabajar los ejercicios y otras tareas; intente ejercicios adicionales para estar seguro que entendió los conceptos. Estudie su desempeño con base en asignaciones, cuestionarios o exámenes previos del semestre.
- Estudie sus apuntes de clases y discusiones. Ponga atención a lo que su instructor espera que usted sepa para un examen.
- Vuelva a leer las secciones relevantes en el libro de texto, poniendo especial atención a notas que haya hecho en los márgenes.
- Estudie de manera individual antes de unirse a un grupo de estudio con amigos. Los grupos de estudio son efectivos sólo si *cada* individuo va preparado para contribuir.
- No llegue demasiado tarde antes de un examen. No coma muchos alimentos una hora antes del examen (pensar es más difícil cuando la sangre está siendo desviada al sistema digestivo).
- Intente relajarse antes y durante el examen. Si ha estudiado de manera efectiva, usted será capaz de hacerlo bien. Permanecer relajado le ayudará a pensar con claridad.

Índice de aplicaciones

(CS = CASO DE ESTUDIO, E = EJEMPLO, H = SECCIÓN HABLEMOS DE..., ET = EJEMPLO EN EL TEXTO, P = PROBLEMA, PR = PROYECTO)

ARTES Y LITERATURA					
PREMIOS DE LA ACADEMIA	P	3.1, p. 96	PRUEBAS DE DROGAS	P	4.4, pp. 184-185
	E	3.2, p. 106	USO DE DROGAS EN ADOLESCENTES	P	9.3, p. 399
IDENTIFICACIÓN DE AUTORES	P	4.3, p. 175	ESPECIES EN PELIGRO DE EXTINCIÓN	PR	3.1, p. 98
SINFONÍAS DE MAHLER Y BEETHOVEN	P	4.3, p. 176	SELECCIÓN DE GÉNERO EN BEBÉS	P	1.3, p. 32
DURACIÓN DE PELÍCULAS	P	5.1, p. 202		P	6.1, p. 237
	PR	5.1, p. 204		ET	9.1, p. 370
SECUELAS DE PELÍCULAS	ET	4.1, p. 146	ALIMENTO GENÉTICAMENTE MODIFICADO	H	CAP. 9, p. 406
DIRECTORES DE ORQUESTA Y LONGEVIDAD	P	7.4, p. 320	GENÉTICA	P	6.5, p. 275
CALIFICACIONES DE LEGIBILIDAD	ET	10.3, p. 429	LESIONES EN LA CABEZA Y ACCIDENTES		
	E	10.3, p. 431	AUTOMOVILÍSTICOS	P	10.3, p. 434
VOCABULARIO DE SHAKESPEARE	H	CAP. 8, p. 366	CHOCOLATE SALUDABLE	P	1.4, p. 42
			ESTILOS DE VIDA SALUDABLES	H	CAP. 1, p. 57
			CARDIOPATÍAS	P	1.3, p. 32
			PULSO CARDIACO	E	5.2, p. 209
			ESTATURA DE HOMBRES	P	5.2, p. 214
			ESTATURA DE MUJERES	E	5.2, p. 210
				P	5.2, p. 213
			LÍNEAS DE ALTO VOLTAJE	P	1.3, p. 32
			RIESGOS DE VIH	P	4.4, p. 185
			ENVENENAMIENTO POR PLOMO	PR	7.3, p. 314
			DURACIÓN DE ESTANCIA EN HOSPITAL	P	9.2, p. 392
			ESPERANZA DE VIDA	E	2.3, p. 70
			ESPERANZA DE VIDA Y MORTALIDAD INFANTIL	E	7.1, p. 290
				ET	7.2, p. 308
			ESPERANZA DE VIDA Y MORTALIDAD	ET	6.4, p. 263
				E	6.4, p. 264
				P	6.4, p. 266
				PR	6.4, p. 266
			ESPERANZA DE VIDA, ESTADOS UNIDOS		
			VS. MUNDO	PR	6.4, p. 266
			PRUEBAS CON LIPITOR	P	9.3, p. 400
			TRATAMIENTOS MAGNÉTICOS	P	1.3, p. 32
			MAMOGRAFÍA	ET	4.4, p. 179
			MELANOMA	P	3.3, p. 122
				E	6.4, p. 262
			NICOTINA Y CIGARROS	P	10.1, p. 417
			SALVADO DE AVENA Y CARDIOPATÍAS	CS	7.2, p. 304
			MADRES GRANDES DE EDAD	P	7.4, p. 320
			FUMADORES PASIVOS Y ACTIVOS	E	4.3, p. 167
				PR	4.3, p. 177
			MUERTES PEATONALES	P	6.5, p. 275
			EFEECTO PLACEBO	ET	1.3, p. 26
			EMBARAZO Y BEBER	P	10.2, p. 427
			DURACIÓN DE EMBARAZO	ET	5.1, p. 196
				E	5.2, p. 207
				P	5.2, p. 213
				P	6.5, p. 277
			PRUEBAS DE EMBARAZO	ET	8.2, p. 346
			INGESTA DE PROTEÍNAS POR HOMBRES	E	8.2, p. 349
			INGESTA DE PROTEÍNAS POR MUJERES	E	1.4, p. 39
			RADÓN Y CÁNCER DE PULMÓN	PR	1.3, p. 33
			MEMORIA REPRIMIDA	CS	1.1, p. 8
			VACUNA CONTRA LA POLIO DE SALK	E	1.3, pp. 23-27
				E	6.1, p. 236
CIENCIAS BIOLÓGICAS Y DE LA SALUD					
ALCOHOL Y CHOQUES	PR	7.4, p. 321			
ASBESTOS	PR	7.3, p. 314			
ASPIRINA Y CARDIOPATÍAS	E	1.4, p. 35			
RED DE CAMA Y MALARIA	P	6.1, p. 237			
CASCOS PARA BICICLETA Y LESIONES	P	10.2, p. 427			
MIGRACIÓN DE AVES	E	3.3, p. 118			
TASAS DE NATALIDAD/MORTALIDAD, CHINA	P	6.4, p. 266			
TASAS DE NATALIDAD/MORTALIDAD, ESTADOS UNIDOS	E	3.2, p. 109			
	P	6.4, p. 266			
TASAS DE NATALIDAD/MORTALIDAD, MUNDIAL	P	7.2, p. 306			
PESO AL NACER	P	9.2, p. 392			
PESO AL NACER Y COMPLEMENTOS VITAMÍNICOS	P	10.1, p. 417			
GRUPOS SANGUÍNEOS	PR	6.2, p. 250			
PRESIÓN SANGUÍNEA	P	5.3, p. 222			
	E	10.1, p. 412			
	P	10.3, p. 434			
ÍNDICE DE MASA CORPORAL	P	1.1, p. 9			
TEMPERATURA CORPORAL	P	1.1, p. 10			
	P	6.1, p. 237			
	E	9.2, p. 387			
CIRUGÍA CARDIACA DE BYPASS	CS	7.4, p. 317			
SÍNDROME DEL TÚNEL CARPIANO	P	6.1, p. 237			
TELÉFONOS CELULARES Y ACCIDENTES	P	10.3, p. 436			
NIÑOS Y BOLSAS DE AIRE	CS	7.4, p. 317			
	PR	7.4, p. 320			
COLESTEROL	E	5.2, p. 210			
CAFÉ Y CÁLCULOS BILIARES	PR	7.4, p. 320			
CONFUSIÓN EN ESTUDIO DE DROGA	CS	1.3, p. 25			
COTININA Y FUMAR	P	3.2, p. 110			
COSTOS HOSPITALARIOS POR ACCIDENTES	P	10.1, p. 416			
TASAS DE MORTALIDAD Y LÍMITES DE VELOCIDAD	P	7.1, p. 296			
TASAS DE MORTALIDAD DEBIDAS A ENFERMEDADES	E	3.3, p. 113			
FIBRA DIETÉTICA	PR	7.4, p. 320			
PRUEBAS DE ENFERMEDADES	P	4.4, p. 184			
HUELLAS DEL DNA	H	CAP. 6, p. 282			
DROGAS EN PELÍCULAS	P	8.3, p. 359			

TAMAÑO DEL PECHO DE MILITARES	
ESCOCES	E 5.1, p. 198
EFFECTOS ESTACIONALES DE ESQUIZOFRENIA	P 3.4, p. 132
CINTURONES DE SEGURIDAD Y NIÑOS	P 6.1, p. 237
TAMAÑO DE CRÁNEOS	P 10.3, p. 434
FUMAR Y CÁNCER	E 1.4, p. 36
	ET 7.4, p. 316
	P 7.4, p. 320
TASAS DE FUMADORES	P 9.3, p. 399
HÁBITOS DE SALUD EN ESTADOS	P 8.3, p. 359
ESTUDIO DE SÍFILIS	PR 1.3, p. 33
TOQUE TERAPÉUTICO	P 1.3, p. 32
	H CAP. 10, p. 443
USO DE TABACO	P 3.4, p. 131
	P 7.4, p. 322
TRATAMIENTO DE ÁLAMOS	P 4.3, p. 175
MUERTES POR TUBERCULOSIS	P 4.4, p. 183
NACIMIENTO DE MELLIZOS	P 3.4, p. 135
ESTUDIOS DE MELLIZOS	PR 1.4, p. 44
MELLIZOS, SUECOS	P 1.3, p. 32
VASECTOMÍA	P 7.4, p. 320
NUTRICIÓN MUNDIAL Y MORTALIDAD	
INFANTIL	P 7.1, p. 298

ADMINISTRACIÓN Y ECONOMÍA

AEROLÍNEAS, A TIEMPO	PR 1.1, p. 11
ÍNDICE DE CONFIANZA DEL CONSUMIDOR	PR 2.4, p. 81
ÍNDICE DE PRECIOS AL CONSUMIDOR	ET 2.4, p. 77
	E 2.4, p. 78
	P 2.4, p. 80
	PR 2.4, p. 81
	P 3.4, p. 131
PRECIOS DE DIAMANTES	ET 7.1, p. 286
ÍNDICE INDUSTRIAL DOW JONES	E 3.1, p. 92
	P 3.2, p. 111
	H CAP. 4, p. 188
TAMAÑO DE GRANJAS, ESTADOS UNIDOS	E 7.1, p. 291
	E 8.1, p. 340
PRESUPUESTO FEDERAL	CS 2.2, p. 64
GASTO FEDERAL	P 3.3, p. 122
PRECIOS DE GASOLINA	ET 2.4, p. 75
	E 2.4, p. 76
GASTO EN CUIDADO DE LA SALUD	P 2.4, p. 82
PRECIOS DE CASAS	P 3.3, p. 121
ÍNDICE DE PRECIOS DE CASAS	P 2.4, pp. 80-81
INGRESO POR GÉNERO	P 3.4, p. 130
DISTRIBUCIÓN DE INGRESOS	H CAP. 4, p. 191
	P 9.2, p. 393
INGRESO DE LOS PADRES	ET 3.3, p. 112
INFLACIÓN Y DESEMPLEO	PR 7.1, p. 298
INVERSIONES	E 3.3, p. 113
	P 4.3, p. 177
BÚSQUEDA DE TRABAJO	P 3.2, p. 111
DEVOLUCIONES (PAGOS) EN SEGUROS DE VIDA	ET 6.3, p. 252
PRODUCCIÓN DE CARNE Y DE GRANOS,	
MUNDIAL	P 7.1, p. 295
SALARIO MÍNIMO	P 2.4, p. 82
	P 3.4, p. 132
UTILIDADES DE PELÍCULAS	E 7.1, p. 293
	P 7.1, p. 297
RATINGS DE NIELSEN	ET 1.1, p. 2
	PR 1.1, p. 10

	E 1.2, p. 13
	P 2.4, p. 80
	E 8.3, p. 356
CONSUMO DE PETRÓLEO, ESTADOS UNIDOS	P 3.4, p. 130
PORCENTAJE DE PROPIETARIOS DE CASA	ET 3.4, p. 126
ÍNDICE DE PRECIOS AL PRODUCTOR	PR 2.4, p. 81
SEGURIDAD SOCIAL	CS 6.4, p. 264
	E 7.1, p. 288
ESTÁNDAR DE VIDA	H CAP. 2, p. 86
PREDICCIONES EN EL MERCADO DE VALORES	E 7.2, p. 303
REDUCCIÓN DE IMPUESTOS	H CAP. 2, p. 84
PRINCIPALES VENDEDORES AL MENUDEO	P 7.1, p. 297
DESEMPLEO	E 1.1, p. 4
	P 1.1, p. 10
	PR 1.1, p. 11
	E 1.2, p. 16
	PR 1.2, p. 20
	E 8.3, p. 356
	P 8.3, p. 360

EDUCACIÓN

COSTO DE LA UNIVERSIDAD	P 2.4, p. 80
	ET 3.4, p. 129
GRADOS UNIVERSITARIOS POR GÉNERO	P 3.3, p. 122
ESTUDIANTES UNIVERSITARIOS POR GÉNERO	P 3.3, p. 120
EDUCACIÓN Y LONGEVIDAD	H CAP. 9, p. 403
GÉNERO Y GRADOS EN ADMINISTRACIÓN/	
BIOLOGÍA	ET 10.2, p. 418
PRESIÓN POR CALIFICACIONES	P 8.3, p. 360
	P 9.3, p. 399
PUNTUACIONES DEL GRE	P 5.2, p. 213
PUNTUACIÓN DEL CI	E 5.2, pp. 209-210
	P 5.2, p. 213
	H CAP. 5, p. 229
	H CAP. 7, p. 325
APRENDIZAJE PARA LEER	
CALIFICACIONES EN EXÁMENES DE NEBRASKA	
Y NJ	P 4.4, p. 183
CALIFICACIONES DEL SAT	PR 5.1, p. 203
	E 5.2, p. 205
	P 5.2, p. 213
	H CAP. 5, p. 226
	P 6.1, p. 238
SEGREGACIÓN ESCOLAR	P 3.3, p. 123
SALARIOS INICIALES DE UNIVERSITARIOS	
GRADUADOS	P 9.3, p. 393
ESPECIALIDADES DE ESTUDIANTES	E 3.2, p. 104
TIEMPO PARA GRADUARSE	P 8.2, p. 354
MUJERES EN EDUCACIÓN SUPERIOR	ET 3.4, p. 127
	P 9.3, p. 399

CIENCIAS FÍSICAS Y AMBIENTALES

AMENAZA DE ASTEROIDE	CS 3.4, p. 127
MILLAJE DE GASOLINA EN AUTOMÓVILES	P 9.3, p. 392
CONTAMINACIÓN DE AUTOMÓVILES	P 10.1, p. 416
EMISIONES DE CO ₂	E 3.2, p. 101
	PR 3.2, p. 111
CONTORNOS DE ELEVACIÓN	E 3.3, p. 117
USO DE ENERGÍA POR ESTADO	ET 3.1, p. 91
	PR 3.1, p. 98
	ET 3.2, p. 105
	ET 3.3, p. 116
CALENTAMIENTO GLOBAL	E 2.2, p. 60

	H	CAP. 3, p. 140
	H	CAP. 7, p. 328
PROBABILIDADES DE HURACÁN	E	6.2, p. 243
PLOMO EN EL AIRE	P	10.1, p. 416
USO DE GAS NATURAL	P	9.3, p. 399
OCEANOS Y MARES	E	4.1, p. 148
ERUPCIONES DEL VIEJO FIEL	P	4.1, p. 154
	P	4.2, p. 162
PLANETAS DE OTRAS ESTRELLAS	E	1.4, p. 3.8
CONFIABILIDAD DE RADIO	P	10.1, p. 416
CONTORNOS DE TEMPERATURA	ET	3.3, p. 117
PRECISIÓN EN PRONÓSTICO DE CLIMA	P	4.3, p. 175
	E	7.1, p. 292
	P	7.1, p. 295
	P	10.1, p. 416
CLIMA MUNDIAL	P	7.1, p. 297
	P	7.2, p. 306

CIENCIAS SOCIALES

ADOPCIÓN DE CHINA	P	3.4, p. 134
EDAD Y GÉNERO	PR	6.2, p. 250
EDAD EN EL PRIMER MATRIMONIO	P	6.2, p. 249
DISTRIBUCIÓN DE EDADES, ESTADOS UNIDOS	E	3.2, p. 107
	P	3.4, p. 131
	PR	6.2, p. 250
POBLACIÓN EN LA EXPLOSIÓN DEMOGRÁFICA	P	3.4, p. 132
CENSO, ESTADOS UNIDOS	CS	2.2, p. 61
ABUSO DE NIÑOS	E	1.3, p. 28
ARMAS OCULTAS	P	8.3, p. 360
CRIMEN Y DESCONOCIDOS	P	10.3, p. 427
NIÑOS AGRESIVOS POR GUARDERÍAS	H	CAP. 1, p. 48
DROGAS PARA DEPRESIÓN	CS	1.3, p. 31
DETECCIÓN DE FRAUDE	H	CAP. 10, p. 439
MANEJAR Y TOMAR	P	3.3, p. 123
	P	3.4, p. 135
EMAIL Y PRIVACIDAD	P	10.2, p. 427
ACCIDENTES MORTALES POR ARMAS DE FUEGO	P	3.3, p. 124
PRODUCCIÓN DE BASURA	E	8.2, p. 350
	P	8.2, p. 354
DECLARACIÓN DE CULPABLE Y PRISIÓN	P	6.5, p. 265
	E	10.2, p. 425
CONTROL DE ARMAS	P	7.4, p. 320
	E	10.2, p. 419
ESTADÍSTICAS DE CONTRATACIÓN	P	4.4, p. 185
HISTORIA DE LA ESTADÍSTICA	H	CAP. 8, p. 363
TASA DE HOMICIDIOS, ESTADOS UNIDOS	ET	3.2, p. 108
ÍNDICE DE DESARROLLO HUMANO	PR	2.4, p. 81
SELECCIÓN DE JURADO	E	6.5, p. 270
NIÑOS Y LOS MEDIOS	PR	4.3, p. 177
POBLACIÓN ZURDA	E	9.3, p. 397
MATRIMONIO Y DIVORCIO	P	3.3, p. 121
ADULTOS CASADOS	P	9.3, p. 399
CENTRO MEDIO DE LA POBLACIÓN,		
ESTADOS UNIDOS	P	4.1, p. 156
MINORÍAS Y POBREZA	E	6.5, p. 273
MUERTES POR VEHÍCULOS MOTORIZADOS	P	3.2, p. 111
MARCHA DE NAPOLEÓN	H	CAP. 3, p. 137
USO DE COMPUTADORAS PERSONALES	E	3.3, p. 114
PRUEBAS DE POLÍGRAFO	P	4.4, p. 184
DENSIDAD DE POBLACIÓN POR ESTADO	E	5.1, p. 198
CRECIMIENTO POBLACIONAL POR ESTADO	P	7.1, p. 296
POBREZA EN IDAHO	P	9.3, p. 399

ENCUESTAS PRE-ELECTORALES	ET	9.3, p. 394
EDADES CUANDO ASUMEN LA PRESIDENCIA	P	4.3, p. 176
PRESIÓN EN ADOLESCENTES	P	8.3, p. 360
PSICOLOGÍA Y RIESGO	PR	6.3, p. 258
SENTENCIAS POR DESFALCO	P	9.2, p. 393
FUMADORES EN CHINA	P	10.3, p. 427
RIESGO EN VIAJE	E	6.4, p. 261
	P	6.4, p. 265
TV E INGRESOS	P	7.1, p. 297
AFLUENCIA A VOTAR Y DESEMPLEO	E	7.3, p. 311
	CS	2.2, p. 64
	PR	2.2, p. 67
MADRES QUE TRABAJAN	P	3.3, p. 124
POBLACIÓN MUNDIAL	E	2.3, p. 69
	PR	2.3, p. 74
	ET	3.4, p. 129
INDICADORES DE POBLACIÓN MUNDIAL	PR	7.3, p. 315

ENCUESTAS Y SONDEOS DE OPINIÓN

ABORTO	P	1.1, p. 10
	P	1.4, p. 43
	P	9.3, p. 399
PERFORACIÓN EN ANWR	P	1.1, p. 9
CONTROL DE NATALIDAD	E	1.4, p. 40
PROPIETARIOS DE AUTOMÓVILES	P	9.2, p. 392
CAMBIO DE CARRERAS	P	1.1, p. 10
CLONACIÓN	E	1.1, p. 6
	P	1.2, p. 19
COMPRADORES COMPULSIVOS	P	9.2, p. 392
¿USTED VOTA?	P	1.1, p. 10
ESTUDIANTES UNIVERSITARIOS DE PRIMER AÑO	P	8.3, p. 359
KENTUCKY FRIED CHICKEN	P	2.4, p. 82
VIDA EN MARTE	E	1.1, p. 6
REVISTA DE LITERATURA	CS	1.4, p. 37
DINERO Y AMOR	E	1.4, p. 38
ALEATORIZACIÓN DE UNA ENCUESTA	PR	6.2, p. 250
PREFERENCIA DE LECTURA	P	3.2, p. 110
CUENTAS DE AHORRO	P	1.1, p. 10
INVESTIGACIÓN DE CÉLULAS MADRE	P	1.1, p. 9
	P	1.3, p. 32
REDUCCIÓN DE IMPUESTOS	E	1.4, p. 40
	P	1.4, p. 43
PRONÓSTICO DE CLIMA,	P	1.3, p. 32
¿QUÉ ES IMPORTANTE?	P	8.3, p. 359

VARIOS (INCLUYENDO DEPORTES)

ANTIGÜEDAD DE AERONAVES	P	5.3, p. 221
TASAS DE ACCIDENTES EN AVIACIÓN	P	6.4, p. 265
PERDICIÓN DE PETE ROSE	E	8.3, p. 357
JUGADOR MÁS VALIOSO EN BÉISBOL	P	7.1, p. 296
SALARIOS EN BÉISBOL	E	2.4, p. 78
REGISTRO DE GANADOS/PERDIDOS		
EN BALONCESTO	P	4.4, p. 184
PROMEDIO DE BATEO	P	4.3, p. 177
DISTANCIA DE FRENADO EN AUTOMÓVIL	P	3.4, p. 130
PRUEBAS DE ACCIDENTES EN AUTOMÓVIL	E	10.1, p. 414
DADOS EN UN CASINO	P	6.3, p. 258
SUSCRIPCIONES DE TELÉFONOS CELULARES	P	3.4, p. 131
ESCÁNER DE VERIFICACIÓN DE COMPRA	E	6.2, p. 244
PESO DE MONEDAS	P	5.2, p. 213
	P	9.2, p. 392

VELOCIDAD DE COMPUTADORAS	ET	3.4, p. 127	CARRERA DE CABALLOS	P	6.3, p. 257
	P	3.4, p. 131	EFFECTO MOZART	E	1.3, p. 25
PERIÓDICOS	P	3.3, p. 123	ESTATURA DE JUGADORES DE LA NBA	P	10.1, p. 415
DISTANCIA DEL CODO A LA PUNTA DE LOS DEDOS	P	10.1, p. 415	ÉXITO EN LA NFL	PR	7.1, p. 298
INFLAMABILIDAD DE TEJIDOS	P	10.3, p. 433	TIEMPOS DE CARRERAS EN OLIMPIADAS	P	10.1, p. 417
PUNTOS EXTRA EN FUTBOL	P	6.3, p. 257	TECLADOS QWERTY	P	3.1, p. 97
FALACIA DEL JUGADOR	ET	6.3, p. 254	MILLAJE DE AUTOMÓVILES EN RENTA	E	9.1, p. 377
	P	6.3, p. 258		E	9.2, p. 385
ATLETAS NOVATOS	P	8.3, p. 359	RULETA	P	6.3, p. 258
VESTIRSE DE NARANJA PARA CAZAR	P	6.5, p. 274	PESOS DE JUGADORES DE RUGBY	P	4.2, p. 162
LOTERÍAS	P	6.2, p. 249	EDADES DE POLIZONES	P	4.3, p. 175
	E	6.3, p. 253	TITANIC	P	2.4, p. 82
	P	6.3, p. 257	DESCARRILAMIENTO DE TRENES	P	3.2, p. 111
	PR	6.3, p. 259	MUJERES EN JUEGOS OLÍMPICOS	E	3.3, p. 119
	H	CAP. 6, p. 279	TIEMPOS DE RÉCORD MUNDIAL EN UNA MILLA	E	7.3, p. 310



Por regla general, quien es más exitoso en la vida es quien posee la mejor información.

—Benjamin Disraeli

Hablemos de estadística

¿EL AGUA QUE TOMA ES SEGURA? ¿LA MAYORÍA DE LAS personas aprueban el plan de impuestos del presidente? ¿Estamos obteniendo un buen valor por el gasto en el cuidado de nuestra salud? Preguntas como éstas sólo pueden abordarse por medio de estudios estadísticos. En el primer capítulo analizamos los principios básicos de la investigación estadística y establecemos el fundamento para un estudio más detallado de la estadística en el resto de esta obra. A lo largo del texto consideraremos diversos ejemplos que muestran cómo estudios estadísticos, bien diseñados, proporcionan una guía para la toma de decisiones en política social y personal, así como algunos casos en los que la estadística puede ser engañosa o mal interpretada.

OBJETIVOS DE APRENDIZAJE

1.1 ¿Qué es la estadística?

Comprender los dos significados del término *estadística* y las ideas básicas que respaldan cualquier investigación estadística, incluyendo las relaciones entre la población de estudio, la muestra, la muestra estadística y los parámetros de la población.

1.2 Muestreo

Entender la importancia de seleccionar una muestra representativa y familiarizarse con los diversos métodos comunes de muestreo.

1.3 Tipos de estudios estadísticos

Entender las diferencias entre los estudios de observación y los experimentos; reconocer las partes importantes en los experimentos, como la selección del tratamiento y los grupos de control, el efecto del placebo y las pruebas ciegas.

1.4 ¿Debe tener confianza en una investigación estadística?

Ser capaz de evaluar estudios estadísticos que oiga en los medios, de modo que pueda decidir si los resultados son significativos.

1.1 ¿Qué es la estadística?

*El pensamiento estadístico
algún día será tan necesario
para la ciudadanía como la
habilidad de leer y escribir.*

—H. G. Wells

Si usted, como la mayoría de los estudiantes que utilizan este texto, es un novato en el estudio de la estadística, podría no estar seguro del porqué estudia estadística y entonces sentir ansiedad acerca de hacia dónde va. Pero conforme empiece a leer, deseamos que su interés crezca y tenga una agradable sorpresa.

El tema de la estadística con frecuencia está estereotipado como árido o técnico, pero incide en todos los ámbitos de la sociedad moderna. Por ejemplo, la estadística nos permite saber si un nuevo medicamento es efectivo en el tratamiento del cáncer, ayuda a los inspectores agrícolas a estar seguros que nuestra alimentación es sana y es la clave de todas las encuestas de opinión. Los negocios utilizan la estadística en investigación de mercado y en publicidad. Hacemos un uso constante de la estadística en los deportes, como una manera de clasificar a los equipos y a los jugadores. En realidad, es difícil pensar en un tema que no esté vinculado con la estadística de algún modo importante.

El objetivo principal de este libro es ayudarlo a aprender las ideas centrales en las que se fundamentan los métodos estadísticos. Estas ideas básicas no son difíciles de entender, aunque adquirir un dominio de los detalles y de la teoría detrás de ellos puede requerir de años de estudio. Una de las grandes ventajas de la estadística es que incluso la poca teoría que aborda este libro le dará la capacidad para entender la estadística que encuentra en las noticias, en sus clases o en el trabajo y en su vida diaria.

Un buen lugar para iniciar es el término *estadística* o *estadísticas*, que si se utiliza en singular o en plural tiene significado diferente. Cuando es singular, *estadística* es la *ciencia* que ayuda a entender cómo recolectar, organizar e interpretar números u otra información acerca de algún tema; haremos mención a los números u otras partes de información como *datos*. Cuando es plural, *estadísticas*, son los datos que describen alguna característica. Por ejemplo, si existen 30 estudiantes en su grupo y sus edades varían de 17 a 64, los números “30 estudiantes”, “17 años” y “64 años” son estadísticas que describen, de alguna manera, a su grupo.

Definiciones de estadística y estadísticas

- Estadística es la *ciencia* que recolecta, organiza e interpreta datos.
- Estadísticas son los *datos* (números y otras partes de información) que describen o resumen algo.

Cómo funciona la estadística

¿Ha visto el Súper Tazón? Los anunciantes lo necesitan saber, ya que el costo del tiempo de publicidad durante el gran juego es aproximadamente de 3 millones de dólares por un *spot* de 30 segundos. Este espacio vale bien su precio si suficientes personas lo observan. Por ejemplo, los reportes de noticias establecen que 93.2 millones de estadounidenses observaron ganar el Súper Tazón XLI a los Potros de Indianápolis. Pero, podría preguntarse: ¿quién contó a todas estas personas?

La respuesta es *nadie*. La afirmación que 93.2 millones de personas observaron el Súper Tazón resulta de investigaciones estadísticas llevadas a cabo por una compañía llamada Nielsen Media Research. Esta compañía publicó los resultados de sus estudios en el famoso *Nielsen ratings*. Lo sorprendente es que Nielsen elaboró estos índices de audiencia por un monitoreo de los hábitos de los telespectadores en sólo 5 000 hogares.

Si usted es nuevo en el estudio de la estadística, la conclusión de Nielsen puede parecer una exageración. ¿Cómo puede alguien sacar una conclusión acerca de millones de personas si sólo estudia a unos cuantos miles? Sin embargo, la ciencia estadística muestra que esta conclusión puede ser muy precisa, siempre y cuando el estudio sea llevado de manera adecuada.

Apropósito...

La estadística tuvo su origen en la recolección de información para censo e impuestos, que son asunto de Estado. Razón por la cual la palabra *Estado* es la raíz de la palabra *estadística*.

Considere los índices de audiencia de Nielsen del Súper Tazón como un ejemplo y hagamos algunas preguntas clave que ilustrarán cómo, en general, funciona la estadística.

¿Cuál es el objetivo de la investigación?

El objetivo de Nielsen es determinar el número total de estadounidenses que vieron el Súper Tazón. En el lenguaje estadístico decimos que Nielsen está interesada en la **población** de los estadounidenses. El número que Nielsen desea determinar —el número de personas que vieron el Súper Tazón— es una característica particular de la población. En estadística, las características de la población se denominan **parámetros de la población**.

Aunque por lo regular consideramos a una población como un grupo de personas, una población estadística puede ser cualquier clase de grupo, personas, animales o cosas. Por ejemplo, en un estudio de seguridad automovilística, la población podría ser *todos los automóviles en la carretera*. De manera análoga, el término *parámetro de la población* puede referirse a cualquier característica de una población. En el caso anterior, los parámetros de la población podrían incluir el número total de automóviles en la carretera en cierto periodo dado, la tasa de accidentes entre automóviles en la carretera, o el rango de pesos de los automóviles en la carretera.

Definiciones

La **población** en un estudio estadístico es el conjunto *completo* de personas u objetos a estudiar.

Los **parámetros de la población** son las características específicas de la población.

EJEMPLO 1 Poblaciones y parámetros de la población

Para cada una de las situaciones siguientes, describa la población que se estudiará e identifique algunos de los parámetros de la población que serían de interés.

- Usted trabaja para Seguros del Agricultor y ha investigado para determinar el monto promedio pagado a víctimas de accidente en automóviles sin bolsas de aire para impactos laterales.
- Ha sido contratado por McDonald's para determinar los pesos de las papas enviadas cada semana para producir las papas fritas.
- Usted es un reportero de negocios que está cubriendo los Laboratorios Genentech, e investiga si su nuevo tratamiento es efectivo contra la leucemia infantil.

Solución

- La población consiste en las personas que han recibido pagos de seguro por accidentes en automóviles que carecen de bolsas de aire para impactos laterales. El parámetro de la población relevante es la media (promedio) del monto pagado a estas personas. (Vea el capítulo 4 para un análisis más detallado de la media y otras medidas de “promedio”).
- La población consiste en todas las papas enviadas cada semana para hacer papas fritas. Los parámetros relevantes de la población incluyen el peso medio de las papas y la variación de los pesos (por ejemplo, ¿la mayoría de ellos son cercanos o lejanos a la media?).
- La población consiste en todos los niños con leucemia. Parámetros importantes de la población son el porcentaje de niños que se recuperaron *sin* el tratamiento nuevo y el porcentaje de niños que se recuperaron con el tratamiento nuevo.

En realidad, ¿qué se obtiene con los estudios?

Si los investigadores en Nielsen fueran todopoderosos podrían determinar el número de personas que ven el Súper Tazón, preguntando a cada estadounidense. Pero nadie puede hacer eso, en cambio, estiman el número de estadounidenses que lo ven a partir del estudio de un grupo relativamente pequeño de personas. En otras palabras, Nielsen aprende acerca de la población de todos los estadounidenses mediante un monitoreo cuidadoso de los hábitos televisivos

A propósito...

Arthur C. Nielsen fundó su compañía e inventó la investigación de mercado en 1923. Él introdujo el índice de radio Nielsen para clasificar los programas de radio en 1942 y en la década de los sesenta amplió sus métodos para la programación de televisión.



de una **muestra** mucho más pequeña de estadounidenses. En concreto, Nielsen usó dispositivos (conocidos como “medidores de personas”) fijados a los televisores en alrededor de 5 000 hogares, de modo que las aproximadamente 13 000 personas que vivían en estos hogares son la muestra de estadounidenses que estudia Nielsen. (Para seguir el rápido cambio de los hábitos de las personas, que ahora incluye programas en internet y por televisión de cable, Nielsen ha incrementado su número de medidores de personas, pero los 5 000 hogares siguen siendo representativos del proceso general de clasificación).

Las medidas individuales que Nielsen recolecta de las personas en los 5 000 hogares constituyen los **datos**. Nielsen recolecta muchos datos —por ejemplo, cuándo y cuánto tiempo está encendido cada televisor en la casa, qué programa está sintonizado y quién ve el programa—. Luego Nielsen consolida esa información en un conjunto de números que caracterizan a la muestra, tal como el porcentaje de televidentes en la muestra que vieron cada programa de televisión o el número total de personas en la muestra que vieron el Súper Tazón. Estos números se denominan **estadísticas muestrales**.

Definiciones

Una **muestra** es un subconjunto de la población del cual se obtienen los datos.

Las medidas u observaciones reales recolectadas de la muestra constituyen los **datos**.

Las **estadísticas muestrales** son características de la muestra encontradas mediante la consolidación o resumen de los datos.

A propósito...

Según la definición del Departamento del Trabajo de Estados Unidos, alguien que no está trabajando no necesariamente está desempleado. Por ejemplo, las mamás y los papás que permanecen en casa no se contabilizan entre los desempleados, a menos que de manera activa estén tratando de buscar un empleo; la gente que buscó trabajo pero se rindió tampoco se contabilizan como desempleados.

EJEMPLO 2 Encuesta de desempleo

El Departamento del Trabajo de Estados Unidos define la *fuerza de trabajo civil* como todas aquellas personas que están empleadas o bien de manera activa buscan un empleo. Cada mes este departamento informa la tasa de desempleo, que es el porcentaje de personas que están activamente buscando un empleo dentro de toda la fuerza de trabajo civil. Con el fin de determinar la tasa de desempleo, este departamento encuesta 60 000 hogares. Para los informes de desempleo, describa

- a. la población b. la muestra c. los datos d. estadísticas muestrales
- e. parámetros de la población

Solución

- a. La *población* es el grupo del que el Departamento del Trabajo quiere aprender, todas las personas que conforman la fuerza de trabajo civil.
- b. La *muestra* consiste en todas las personas de los 60 000 hogares encuestados.
- c. Los *datos* consisten en toda la información recolectada en la encuesta.
- d. Las *estadísticas muestrales* resumen los datos de la muestra. En este caso, la estadística muestral relevante es el porcentaje de personas en la muestra que están buscando de manera activa trabajo. (El Departamento del Trabajo también calcula una estadística muestral similar para subgrupos en la población, tal como los porcentajes de adolescentes, hombres, mujeres y ex combatientes que están desempleados).
- e. Los *parámetros de la población* son las características de toda la población que corresponden a las estadísticas muestrales. En este caso, el parámetro poblacional relevante es la tasa real de desempleo. Observe que el Departamento del Trabajo en realidad *no* mide este parámetro de la población, ya que los datos se recolectan sólo para la muestra y luego se utilizan para estimar el parámetro de la población.

¿Cómo se relacionan las estadísticas muestrales con los parámetros de la población?

Suponga que Nielsen determina que 31% de las personas en los 5 000 hogares de su muestra observaron el Súper Tazón. Este “31%” es una estadística muestral, ya que caracteriza a la

muestra. Pero lo que en realidad Nielsen quiere conocer es el parámetro poblacional correspondiente, que es el porcentaje de estadounidenses que vieron el Súper Tazón.

No hay forma para que los investigadores de Nielsen conozcan el valor exacto del parámetro de la población, ya que ellos sólo estudiaron una muestra. Sin embargo, los investigadores de Nielsen desean hacer su trabajo de modo que la estadística muestral sea una buena estimación del parámetro de la población. En otras palabras, les gustaría concluir que puesto que 31% de la muestra vio el Súper Tazón, entonces aproximadamente 31% de la población también lo vio. Uno de los propósitos principales de la estadística es ayudar a los investigadores a evaluar la validez de este tipo de conclusiones.

UN MOMENTO DE REFLEXIÓN

Suponga que Nielsen concluye que 30% de los estadounidenses vieron el Súper Tazón. ¿Cuántas personas representan esto? (La población de Estados Unidos es aproximadamente de 300 millones).

La ciencia estadística proporciona métodos que permiten a los investigadores determinar cuán bien una muestra estadística estima un parámetro de la población. Por ejemplo, los resultados de encuestas o sondeos de opinión por lo regular se reportan con algo denominado **margen de error**. Al sumar y restar el margen de error de la estadística muestral, determinamos un rango de valores o **intervalo de confianza**, que es *probable* que contenga el parámetro de la población. En la mayoría de los casos, el margen de error se define de modo que podamos tener 95% de confianza que este rango contiene al parámetro de la población. Analizaremos el significado preciso de “probable” y “95% de confianza” en el capítulo 8, pero por ahora podría serle útil una explicación dada por el *New York Times* (figura 1.1). En el caso de los índices de audiencia de Nielsen, el margen de error es alrededor de un punto porcentual. Por tanto, si 31% de la muestra estuvo viendo el Súper Tazón, entonces podemos tener 95% de confianza que el rango de 30% a 32% contiene al porcentaje real de la población que vio el Súper Tazón.

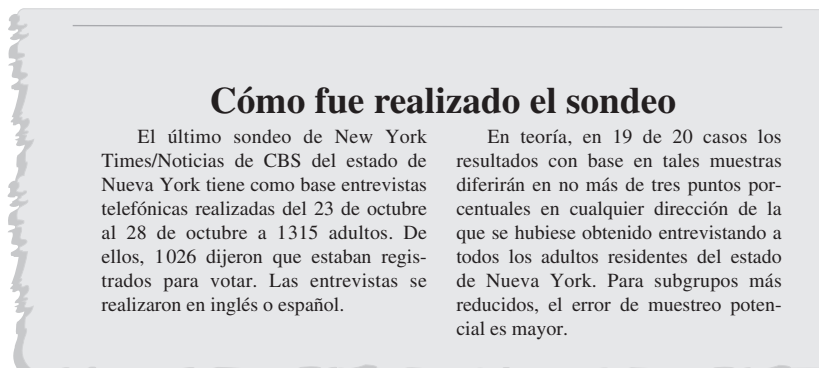


FIGURA 1.1 El margen de error en una encuesta o sondeo de opinión por lo regular describe un rango que es probable (con 95% de confianza, lo que quiere decir 19 de 20 casos) contenga al parámetro de la población. Este extracto del *New York Times* explica un margen de error de 3 puntos porcentuales.

Uno de los hallazgos más extraordinarios de la ciencia estadística es que es posible obtener resultados significativos a partir de muestras pequeñas. Sin embargo, muestras de tamaño grande son mejores (cuando son factibles), ya que el margen de error, por lo común, es menor para muestras mayores. Por ejemplo, el margen de error para un intervalo de confianza de 95% en una encuesta bien realizada es alrededor de 5 puntos porcentuales para un tamaño de muestra de 400, pero baja a 3 puntos porcentuales para un tamaño de muestra de 1 000 y a 1 punto porcentual para una muestra de 10 000. (Vea el capítulo 8 para entender cómo se calculan los márgenes de error).

Definición

El **margen de error** en una investigación estadística sirve para describir el rango de valores, o **intervalo de confianza**, que probablemente contenga al parámetro de la población. Este intervalo de confianza lo determinamos al sumar y restar el margen de error a la estadística muestral obtenida en el estudio. Es decir, el rango de valores que probablemente contengan el parámetro de la población es

desde (estadística muestral – margen de error)

hasta (estadística muestral + margen de error)

Por lo común el margen de error es definido para que proporcione un intervalo de 95% de confianza, esto quiere decir que 95% de las muestras del tamaño utilizado en el estudio contendrían el parámetro real de la población (y 5% no lo contendrían).

EJEMPLO 3 ¿Vida en Marte?

El Centro de Investigación Pew para la Gente y la Prensa entrevistó a 1 546 adultos estadounidenses acerca de sus actitudes hacia el futuro. Les preguntó si el hombre llegaría a Marte en los próximos 50 años, 76% de estas 1 546 personas dijeron *definitivamente sí* o *probablemente sí*. El margen de error para la encuesta fue de 3 puntos porcentuales. Describa la población y la muestra para esta encuesta, y explique el significado del estadístico muestral de 76%. ¿Qué puede concluir acerca del porcentaje de la población que piensa que el hombre llegará a Marte en los próximos 50 años?

Solución La población son todos los adultos estadounidenses y la muestra consiste en las 1 546 personas que fueron entrevistadas. El estadístico muestral de 76% es el porcentaje *real* de personas en la muestra que respondió que el hombre definitiva o probablemente llegaría a Marte en los próximos 50 años. El estadístico muestral de 76% y el margen de error de 3 puntos porcentuales nos dicen que el rango de valores

desde $76\% - 3\% = 73\%$

hasta $76\% + 3\% = 79\%$

es probable (con 95% de confianza) que contenga el parámetro de la población, que en este caso es el porcentaje verdadero de todos los adultos estadounidenses que piensan que el hombre definitiva o probablemente llegará a pisar Marte en los próximos 50 años.

EJEMPLO 4 ¿Clonación de humanos?

La misma encuesta Pew preguntó a las personas si creían que los humanos serían clonados en los siguientes 50 años. A esta pregunta 51% respondió *definitivamente sí* o *probablemente sí*. De nueva cuenta, el margen de error fue de 3 puntos porcentuales. ¿Podemos estar seguros que una mayoría de los adultos estadounidenses piensan que los humanos serán clonados en los próximos 50 años?

Solución No. Para determinar el rango de valores que probablemente contenga el porcentaje de *todos* los adultos estadounidenses quienes piensan que la clonación humana definitiva o probablemente ocurra, sumamos y restamos el margen de error de 3 puntos porcentuales de la estadística muestral de 51%. Esto proporciona un rango de valores de 48% a 54%. Puesto que este rango incluye valores a ambos lados de 50%, no podemos estar seguros que la mayoría (es decir, más de 50%) de los adultos estadounidenses piense que los humanos serán clonados en los siguientes 50 años.

UN MOMENTO DE REFLEXIÓN

Busque un reporte sobre una encuesta de opinión en las noticias de esta semana. ¿El reporte proporciona un margen de error? En este caso, ¿qué significa?

A propósito...

Un clon es una copia genética exacta de su padre. Por ejemplo, un clon de usted sería genéticamente igual a usted, pero nacería como un bebé y tendría experiencias de vida distintas a las suyas. La primera clonación exitosa de un mamífero adulto fue en 1997, cuando Ian Wilmut y sus colegas en Escocia clonaron a una oveja. El clon, llamado Dolly, no tiene padre, ya que *todos* sus genes provienen de la madre. En la actualidad los científicos han clonado a muchos otros mamíferos.



Cómo reunir las partes: el proceso de una investigación estadística

El proceso utilizado por Nielsen Media Research es similar al que se utiliza en muchos estudios estadísticos. La figura 1.2 y el recuadro siguiente resumen los pasos básicos en una investigación estadística. Tenga en cuenta que estos pasos están un poco idealizados, y los pasos reales pueden diferir de un estudio a otro. Además, los detalles ocultos en los pasos básicos son muy importantes. Por ejemplo, una mala elección de la muestra en el paso 2 puede causar que todo el estudio carezca de significado, y debe tenerse sumo cuidado en inferir conclusiones acerca de una población tomando como base resultados de una muestra muy pequeña de la población.

Pasos básicos en una investigación estadística

- Paso 1. Indique con precisión el objetivo de su estudio; esto es, determine la población que quiere estudiar y lo que quiere saber de ella.
- Paso 2. Seleccione una muestra de la población. (Asegúrese de utilizar una técnica apropiada de muestreo, como se estudia en la sección siguiente).
- Paso 3. Recolecte datos de la muestra y resuma esos datos determinando las estadísticas muestrales de interés.
- Paso 4. Utilice las estadísticas muestrales para realizar inferencias acerca de la población.
- Paso 5. Obtenga conclusiones, determine lo que aprendió y si alcanzó su objetivo.

A propósito...

Con frecuencia los estadísticos dividen su materia en dos ramas principales: **estadística descriptiva**, que trata la *descripción* de los datos por medio de gráficas y estadísticas muestrales, y **estadística inferencial**, que trata con la *inferencia* (o estimación) de los parámetros de la población a partir de los datos muestrales. En este texto, los capítulos 2 al 5 tratan, principalmente, la estadística descriptiva, mientras que los capítulos 6 a 10 están enfocados en la estadística inferencial.

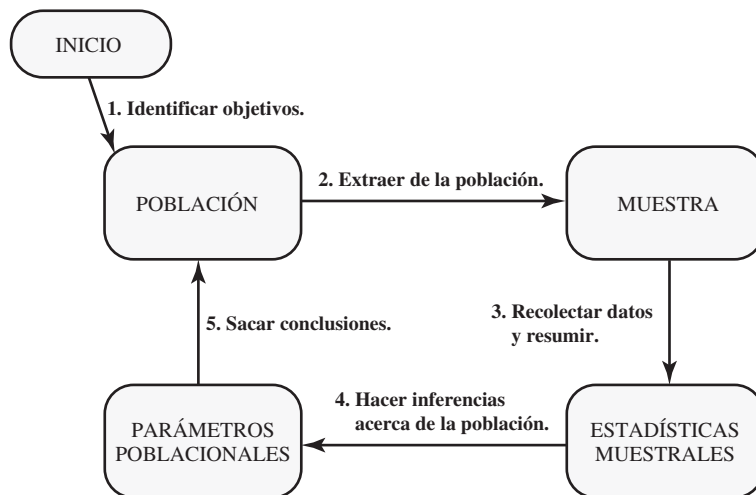


FIGURA 1.2 El proceso de una investigación estadística.

EJEMPLO 5 Identificación de los pasos

Considere la encuesta del Centro de Investigación Pew que se describió en los ejemplos 3 y 4. Ahora identifique cómo los investigadores aplican los cinco pasos básicos en una investigación estadística.

Solución Los pasos se aplican como se indica a continuación.

1. Los investigadores tenían como objetivo aprender acerca de las actitudes específicas que los estadounidenses tienen sobre el futuro. Ellos seleccionaron a los adultos estadounidenses como la población, dejaron intencionalmente fuera a los niños.
2. Seleccionaron 1 546 adultos estadounidenses como su muestra. Aunque no dijeron cómo tomaron la muestra, supondremos que fueron elegidos de modo que los 1 546 adultos estadounidenses son representativos de toda la población de adultos estadounidenses.

3. Ellos recolectaron los datos pidiendo a las personas en la muestra respuestas a las preguntas cuidadosamente seleccionadas. Los datos son las respuestas individuales a las preguntas. Ellos resumieron estos datos con estadísticas muestrales, tal como los porcentajes globales de personas, en la muestra, que respondieron *sí* o *no* a cada pregunta.
4. Las técnicas de la ciencia estadística permitieron a los investigadores inferir características de la población. En este caso, la inferencia consistió en la estimación de ciertos parámetros de la población y el cálculo de los márgenes de error.
5. Seguros de que el estudio fue realizado adecuadamente e interpretando las estimaciones de los parámetros de la población, los investigadores sacaron conclusiones globales sobre las actitudes de los estadounidenses acerca del futuro.

Estadística: decisiones para un mundo incierto

Los índices de audiencia de Nielsen y la mayoría de los ejemplos que hemos analizado hasta ahora incluyen encuestas o sondeos. No obstante, el tema de la estadística abarca mucho más, como diseño de experimentos para probar nuevos tratamientos médicos, análisis de los peligros del calentamiento global e incluso la evaluación del valor de una educación académica. En realidad, es justo decir que el propósito principal de la estadística es ayudar a tomar buenas decisiones siempre que nos enfrentemos con una diversidad de opciones posibles.

Propósito de la estadística

La estadística tiene muchos usos, pero quizás el más importante es ayudarnos a tomar buenas decisiones en situaciones que incluyen incertidumbre.

A propósito...

La poliomielitis se ha vuelto rara en Estados Unidos a partir del desarrollo de la vacuna de Salk, pero sigue siendo común en países menos desarrollados. Un esfuerzo global para vacunar a niños contra la polio inició en 1998 y ha tenido un gran éxito, aunque no se ha alcanzado la meta de erradicar por completo la enfermedad.



La gran recompensa de las obras es la oportunidad de hacer más.

—Jonas Salk

Este propósito será claro en la mayoría de los casos de estudio y ejemplos que consideramos en este texto, pero a veces tendremos que estudiar un poco de teoría que, al principio, podría parecer abstracta. Si tiene en mente el propósito general de la estadística, será recompensado al final, cuando vea cómo la teoría nos ayuda a entender nuestro mundo. El estudio de caso siguiente le dará una muestra de lo que viene. Incluye varias ideas teóricas importantes que llevan a uno de los más grandes logros del siglo XX en salud pública.

ESTUDIO DE CASO

La vacuna de Salk contra la poliomielitis

Si hubiese sido padre en las décadas de los cuarenta o cincuenta, uno de sus mayores temores hubiese sido la enfermedad conocida como poliomielitis. Cada año, durante esta larga epidemia de polio, miles de niños sufrieron parálisis por la enfermedad. En 1954 se realizó un experimento para probar la efectividad de una vacuna nueva creada por el doctor Jonas Salk (1914-1995). El experimento incluyó una muestra de 400 000 niños elegidos de la población infantil en Estados Unidos. La mitad de estos niños recibieron una inyección de la vacuna de Salk. La otra mitad recibió una inyección que sólo contenía agua con sal. (La inyección de agua con sal fue un *placebo*, vea la sección 1.3). Entre los niños que recibieron la vacuna de Salk, sólo 33 contrajeron la polio. En contraste, hubo 115 casos de polio entre los niños que no recibieron la vacuna de Salk. Al utilizar técnicas de la ciencia estadística, que estudiaremos posteriormente, los investigadores concluyeron que la vacuna era efectiva en la prevención de la polio. Por tanto, decidieron emprender un esfuerzo mayor para mejorar la vacuna de Salk y distribuirla a la población de *todos* los niños. Una consecuencia de esto fue que los niños en Estados Unidos y en otros países desarrollados empezaron a recibir la vacuna de manera rutinaria y ahora el horror de la polio es cosa del pasado.

Sección 1.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Población y muestra.** ¿Qué es una población?, ¿qué es una muestra? y ¿cuál es la diferencia entre ellas?
- 2. Estadística y estadísticas.** Suponga que en una plática una persona se refiere a las estadísticas del béisbol, mientras que otra se refiere al uso de la estadística para mostrar que cierto medicamento es un tratamiento efectivo. ¿Ambos usos del término *estadística* tienen el mismo significado? Si no, ¿en qué difieren?
- 3. Estadísticas y parámetros.** ¿Qué es una muestra estadística?, ¿qué es un parámetro de la población? y ¿cuál es la diferencia entre ellos?
- 4. Margen de error.** ¿Qué es el margen de error en una investigación estadística y por qué es importante?

¿Tiene sentido? Para los ejercicios 5 al 10 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Estadísticas y parámetros.** Mi profesor realizó una investigación estadística en la que fue incapaz de medir estadísticas muestrales, pero tuvo éxito en la determinación de parámetros de la población con un muy pequeño margen de error.
- 6. Encuesta fallida.** Un sondeo de opinión llevado a cabo dos semanas antes de las elecciones determinó que Smith obtendría 70% de los votos, con un margen de error de 3%, pero terminó perdiendo la elección.
- 7. Encuesta cierta.** No existe duda que Johnson ganó la elección, ya que una encuesta de salida mostró que ella recibió 54% de los votos y el margen de error fue de sólo 3 puntos porcentuales.
- 8. Venciendo a Nielsen.** Una compañía nueva intenta competir con Nielsen Media Research proporcionando información con un mayor margen de error por el mismo precio.
- 9. Muestra de depresión.** El objetivo de mi estudio es aprender acerca de la depresión en personas que han sufrido una tragedia familiar, por lo que planeo seleccionar una muestra de la población que ha estado enferma en el mes pasado.
- 10. Producto nuevo.** Nuestro departamento de investigación de mercado encuestó a 1 000 consumidores sobre sus actitudes hacia nuestro nuevo producto. Puesto que las personas en la muestra fueron muy entusiastas en su interés por la compra del producto, decidimos extender una campaña de publicidad a nivel nacional.

Conceptos y aplicaciones

Muestra, población, estadística y parámetro. Cada uno de los ejercicios 11 al 14 describe una investigación estadística. En cada caso identifique la muestra, la población, la estadística muestral y el parámetro de la población.

- 11. Investigación de células madre.** En un sondeo del *Newsweek*, realizado por Princeton Survey Research Associates International, se preguntó a 1 002 adultos en Estados Unidos si estaban a favor o en contra del uso de “dinero de impuestos federales para financiar investigación médica que utilice células madre de embriones humanos”. Entre los encuestados, 48% dijeron estar a favor.
- 12. Viernes 13.** En una encuesta de Gallup a 1 236 adultos, 9% respondió que había mala suerte el viernes que fuese día 13 de un mes.
- 13. Distancias galácticas.** Los astrónomos determinan la distancia a una galaxia lejana midiendo las distancias de unas cuantas estrellas en ella y tomando la media de las medidas de estas distancias.
- 14. Medicamento para alergia.** Nasonex es un medicamento utilizado para tratar síntomas de alergia. En una prueba a 374 niños de 3 a 11 años se les dio una dosis de 100 microgramos de Nasonex y 17% de ellos experimentaron dolores de cabeza.

Identificación del rango de valores. En los ejercicios 15 al 18 utilice la estadística dada y el margen de error para identificar el rango de valores (intervalo de confianza) que probablemente contenga al valor verdadero del parámetro de la población.

- 15. Investigación de células madre.** En un sondeo del *Newsweek* realizado por Princeton Survey Research Associates International, se preguntó a 1 002 adultos en Estados Unidos si estaban a favor o en contra del uso de “dinero de impuestos federales para financiar investigación médica que utilice células madre de embriones humanos”. Con base en los resultados del sondeo, 48% de los adultos están a favor, con un margen de error de 3 puntos porcentuales.
- 16. Petróleo en ANWR.** En una encuesta de CBS News/*New York Times* a 1 241 adultos en Estados Unidos se les preguntó si aprobaban o no perforaciones para petróleo y gas natural en el Refugio Nacional de Vida Salvaje del Ártico en Alaska (ANWR, por sus siglas en inglés). Con base en los resultados de la encuesta, 45% de los adultos no la aprueban, con un margen de error de 3 puntos porcentuales.
- 17. Índice de Masa Corporal (IMC).** En Estados Unidos, 40 adultos fueron elegidos de manera aleatoria para medir su IMC. El resultado de esa muestra mostró que el IMC promedio (medio) para hombres es 26.0, con un margen de error de 3.4.

18. Temperatura corporal. Investigadores seleccionaron de manera aleatoria a 106 adultos y midieron la temperatura de su cuerpo. Con base en esa muestra, estimaron que la temperatura promedio (media) del cuerpo es 98.2°F, con un margen de error de 0.1°F.

19. Interpretación de resultados de una encuesta. Una encuesta realizada un día antes de la elección estatal para senadores, en la cual sólo quedan dos candidatos en competencia, muestra que 58% de los votantes están a favor del candidato republicano, con un margen de error de 3 puntos porcentuales. ¿Los republicanos deben esperar ganar? ¿Por qué sí o por qué no?

20. Interpretación de resultados de degustaciones. En una prueba de catar televisada en vivo a nivel nacional, una muestra de bebedores regulares de Budweiser probaron a ciegas muestras de Michelob y Schlitz. La estimación con base en los resultados es que 52% de tales personas tenían preferencia por Michelob, con un margen de error de 10%. ¿Es posible concluir que la mayoría de los tomadores de Budweiser prefieren Michelob, cuando se les da a probar Michelob y Schlitz? ¿Por qué sí o por qué no?

21. ¿Miente la gente acerca de si votó? En una encuesta de 1 002 personas, 701 (o 70%) dijeron que votaron en una elección particular para la presidencia (con base en información de ICR Research Group). El margen de error para esta encuesta fue de 3 puntos porcentuales. Sin embargo, los registros de las votaciones muestran que sólo 61% de todos los votantes elegibles lo hicieron. ¿Esto implica que la gente mintió cuando respondió la encuesta? Explique.

22. ¿Por qué la discrepancia? Una encuesta del Instituto Eagleton preguntó a hombres si estaban de acuerdo con esta afirmación: “El aborto es una cosa privada que debe dejarse a la mujer que decida sin intervención del gobierno”. Entre los hombres que fueron encuestados por mujeres, 77% estaban de acuerdo con la afirmación. Entre los hombres que eran encuestados por hombres, 70% estaban de acuerdo con la afirmación. Suponiendo que la discrepancia es significativa, ¿cómo puede explicar esta discrepancia?

Interpretación de estudios reales. Para cada uno de los ejercicios del 23 al 26, realice lo siguiente:

- Con base en la información dada, indique cuál considera que fue el objetivo del estudio. Identifique una posible población y el parámetro poblacional de interés.
- Describa brevemente la muestra, los datos y la estadística muestral para el estudio.
- Con base en la estadística muestral y el margen de error, identifique el rango de valores (intervalo de confianza) que es probable que contenga al parámetro poblacional de interés.

23. Cambio de carrera. En una encuesta de Korn/Ferry International a 1 733 ejecutivos, 51% de ellos dijeron que si pudiesen iniciar otra carrera nuevamente, elegirían un área diferente. El margen de error fue de 3 puntos porcentuales.

24. Súper poderes. En una encuesta de Caravan a 1 018 adultos, 11% dijeron que elegirían la capacidad de ser invisibles como el súper poder que preferirían. El margen de error fue de 3.2 puntos porcentuales.

25. Tasa de desempleo. Con base en una encuesta reciente a adultos en 60 000 hogares, el Departamento del Trabajo de Estados Unidos reportó una tasa de desempleo de 4.6%. El margen de error fue de 0.2 puntos porcentuales.

26. Ahorro para imprevistos. Una encuesta de la Organización Roper a 2 000 adultos en Estados Unidos mostró que 64% de los encuestados tuvo dinero en cuentas de ahorro regulares. El margen de error fue de 2.0 puntos porcentuales.

Cinco pasos en un estudio. Describa cómo aplicaría los cinco pasos básicos en una investigación estadística (como se listó en el recuadro de la página 7) a los temas en los ejercicios 27 a 30.

27. Teléfonos celulares y manejo. Usted quiere determinar el porcentaje de conductores con licencia que utilizaron un teléfono celular al menos una vez mientras conducían durante la semana pasada.

28. Calificaciones de crédito. La compañía FICO utiliza sus calificaciones para clasificar la calidad de un crédito al consumidor. Usted quiere determinar la calificación promedio (media) de FICO de todos los adultos en Estados Unidos.

29. Peso de pasajeros. Reconociendo que el sobrepeso en una aeronave comercial llevaría a vuelos inseguros, quiere determinar el peso promedio (medio) de los pasajeros en la aerolínea.

30. Pilas de marcapasos. Puesto que las pilas utilizadas en los marcapasos del corazón son sumamente importantes, quiere determinar la duración promedio (media) de tales pilas, a partir de la última falla.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 1 en www.aw.com/bbt.

31. Índices de audiencia actuales de Nielsen. Determine los índices de audiencia de Nielsen para la semana pasada. ¿Cuáles fueron los tres programas de televisión más populares? Explique el significado de “índices de audiencia” y el término “participación” para cada programa.

32. Método Nielsen. La compañía Nielsen Media Research con frecuencia revisa los detalles de sus métodos para reco-

lectar la información. Visite su sitio web y lea acerca de sus estrategias actuales para clasificar los programas de televisión. Resuma estos métodos en una lista con formato de viñetas.

- 33. Comparación de aerolíneas.** El Departamento de Transporte de Estados Unidos publica el desempeño de puntualidad, tasa de equipaje perdido y otras estadísticas para diferentes compañías aéreas. Determine un ejemplo reciente de tales estadísticas. Con base en sus hallazgos, ¿es justo decir que alguna línea en particular destaca como mejor o peor que las otras? Explique.
- 34. Estadísticas laborales.** Utilice el sitio web de la Oficina de Estadísticas de Trabajo para determinar las tasas de desempleo durante los últimos 12 meses. Si supone que la encuesta mensual tiene un margen de error de alrededor de 0.2 puntos porcentuales, ¿ha habido un cambio notable en la tasa de desempleo durante el último año? Explique.
- 35. Estadísticas y seguridad.** Identifique un estudio que se haya realizado (o deba hacerse) para mejorar la seguridad de los automovilistas y los pasajeros. Describa brevemente la importancia de la estadística en el estudio.
- 36. Centro de Investigación Pew.** El Centro Pew para la Gente y la Prensa, estudia actitudes del público hacia la prensa, la política y temas de política. Visite su sitio web y encuentre las últimas encuestas acerca de las actitudes. Seleccione una encuesta, escriba un resumen de lo que fue esa encuesta, cómo se realizó y cuando fue hecha.

EN LAS NOTICIAS

- 37. Estadísticas en las noticias.** Identifique tres artículos de la semana pasada que incluyan estadísticas sobre algo. En cada caso escriba un párrafo breve que describa el papel de la estadística en el artículo.
- 38. Estadística en su carrera.** Escriba una descripción breve de algunas formas en las que piensa que la ciencia estadística puede utilizarse en su campo principal de estudio. (Si no ha seleccionado un área, responda la pregunta para una que esté considerando).
- 39. Estadística y entretenimiento.** Los índices de audiencia de Nielsen son bien conocidos por su papel en el cálculo de televidentes. Identifique otra forma en que las estadísticas son utilizadas en la industria del entretenimiento. Describa brevemente el papel de la estadística en esta aplicación.
- 40. Estadística y deporte.** Seleccione un deporte y describa, al menos, tres estadísticas diferentes que regularmente son seguidas por los participantes o los espectadores. En cada caso, describa la importancia de la estadística para el deporte.
- 41. Estadísticas económicas.** Con regularidad, el gobierno publica muchas estadísticas económicas, tal como la tasa de desempleo, la tasa de inflación y el superávit o déficit en el presupuesto federal. Analice periódicos recientes e identifique cinco estadísticas económicas importantes. Explique brevemente el propósito de cada una de estas cinco estadísticas.

1.2 Muestreo

La única forma de conocer el valor verdadero de un parámetro de la población es observar a *todos* los miembros de la población. Por ejemplo, para aprender acerca de la altura media exacta de todos los estudiantes en su escuela, necesita medir la altura de cada uno de los estudiantes. Una colección de datos de cada miembro de una población se denomina **censo**. Por desgracia, llevar a cabo un censo con frecuencia es poco práctico. En algunos casos la población es tan grande que sería demasiado oneroso o tardaría mucho tiempo recolectar la información de cada miembro. En otros casos un censo interferiría con las metas globales del estudio. Por ejemplo, un estudio diseñado para probar la calidad de las barras de dulce antes de enviarlas a su distribución, no podría incluir un censo ya que significaría probar cada una de las barras de dulce, dejando nada para la venta.

No todo lo que puede contarse cuenta y no todo lo que cuenta puede ser contado.

—Albert Einstein

Definición

Un **censo** es una recolección de datos de *cada uno* de los miembros de una población.

Por fortuna, la mayoría de los estudios estadísticos pueden hacerse sin llevar el problema de la realización de un censo. En lugar de recolectar los datos de cada miembro de la población,

recolectamos la información de una muestra y utilizamos las estadísticas muestrales para hacer inferencias acerca de la población. Por supuesto, las inferencias serán razonables sólo si los miembros de la muestra representan bien a la población, al menos en términos de las características bajo estudio. Esto es, buscamos una **muestra representativa** de la población.

Definición

Una **muestra representativa** es una muestra en la cual las características relevantes de los elementos de la muestra generalmente son las mismas que las características de la población.

EJEMPLO 1 Una muestra representativa para alturas

Suponga que quiere determinar la altura media de todos los estudiantes de su escuela. ¿Cuál podría ser una muestra más representativa para este estudio: los integrantes del equipo varonil de básquetbol o los estudiantes de su curso de estadística?

Solución Los integrantes del equipo varonil de básquetbol no son una muestra representativa para un estudio de altura, ya que sólo está formada por hombres y, segundo, los jugadores de básquetbol tienden a ser más altos que el promedio. La altura media de los estudiantes en su curso de estadística, tal vez, es mucho más cercana a la altura media de todos los estudiantes, de modo que su clase forma una muestra más representativa que los integrantes del equipo varonil de básquetbol.

Sesgo

Imagine que para los 5 000 hogares en su muestra, Nielsen eligió sólo los hogares en los que el principal proveedor de ingresos trabajó en horario nocturno. Puesto que los trabajadores de turno nocturno no están en casa para ver la televisión por la noche, Nielsen determinaría que los programas nocturnos no son populares entre los hogares de esta muestra. Claro, esta muestra *no* sería representativa de todos los hogares de Estados Unidos, sería erróneo concluir que los programas nocturnos no son populares para todos los estadounidenses. Decimos que tal muestra está *sesgada* ya que los hogares en la muestra difieren de una manera específica de los hogares “típicos” estadounidenses. (En realidad, Nielsen tiene gran cuidado en evitar tal sesgo obvio en la selección de la muestra). El término **sesgo** refiere cualquier problema en el diseño o realización de una investigación estadística que tiende a inclinarse a ciertos resultados. No podemos tener confianza de las conclusiones de un estudio sesgado.

Definición

Una investigación estadística sufre de **sesgo** si su diseño o realización tiende a favorecer ciertos resultados.

El sesgo puede surgir de muchas formas. Por ejemplo:

- Una muestra tiene sesgo si los miembros de la muestra difieren de alguna manera específica de los miembros de la población general. En tal caso, los resultados del estudio reflejarán las características inusuales de la muestra en lugar de las características reales de la población.
- Un investigador está sesgado si tiene un interés en un resultado particular. En ese caso, el investigador podría, a propósito o sin querer, distorsionar el significado real de la información.
- El conjunto de datos podría estar sesgado si sus valores fueron recolectados, a propósito o sin querer, de una forma que los hace no representativos de la población.
- Aun si un estudio está bien hecho, podría reportarse de una forma sesgada. Por ejemplo, una gráfica que representa los datos podría decir sólo parte de la historia o ilustrar la información de una manera engañosa (vea la sección 3.4).

A propósito...

Muchos estudios médicos son experimentos diseñados para probar si un medicamento nuevo es efectivo. En un artículo publicado en la *Journal of the American Medical Association*, los autores determinaron que los estudios con resultados positivos (el medicamento es efectivo) es más probable que se publiquen, que los estudios con resultados negativos (el medicamento no es efectivo). Este “sesgo en la publicación” para la producción de nuevos medicamentos, como grupo, parece más efectivo de lo que en realidad son.



Prevenir el sesgo es uno de los mayores retos en una investigación estadística. Por tanto, identificar sesgo es uno de los pasos más importantes en la evaluación de una investigación estadística o de informes en medios acerca de una investigación estadística.

EJEMPLO 2 ¿Por qué usar Nielsen?

Nielsen Media Research cobra a las redes y estaciones de televisiones por sus servicios. Por ejemplo, NBC paga a Nielsen por proporcionar índices de audiencia de sus programas de televisión. ¿Por qué NBC no hace sus propios índices de audiencia, en lugar de pagar a una compañía por hacerlos?

Solución El costo de publicidad en un programa de televisión depende del índice de audiencia del show. Entre mayor sea, más puede cobrar la red por un anuncio, lo que significa que NBC podría tener un claro sesgo si hace sus propios índices de audiencia. Por tanto, los anunciantes no confiarían en estos índices. Al contratar una fuente independiente, tal como Nielsen, NBC proporciona información más creíble para los anunciantes.

UN MOMENTO DE REFLEXIÓN

El hecho que NBC pague a Nielsen por sus servicios podría dar a Nielsen un incentivo financiero para hacer que NBC salga bien en los índices de audiencia. El hecho de que Nielsen también proporciona índices de audiencia para otras cadenas, ¿cómo ayuda a prevenir de un posible sesgo hacia NBC?

Métodos de muestreo

Una buena investigación estadística *debe* tener una muestra representativa. De otra forma la muestra está sesgada y las conclusiones del estudio carecen de confiabilidad. Examinaremos unos cuantos métodos comunes de muestreo que, al menos en principio, pueden proporcionar una muestra representativa.

Muestras aleatorias simples

En la mayoría de los casos, la mejor forma de obtener una muestra representativa es seleccionándola *al azar* (o *aleatoriamente*) de la población. Una **muestra aleatoria** es aquella en la cual cada elemento de la población tiene igual oportunidad de ser seleccionado como parte de la muestra. Por ejemplo, podría obtener una muestra aleatoria que contenga a todos aquéllos de una población que tiren un dado y seleccionen a los que obtengan un 6. En contraste, la muestra no sería aleatoria si selecciona a aquellos de más de 6 pies de estatura, ya que no todos tendrían igual oportunidad de ser seleccionados.

En estadística, por lo regular decidimos por adelantado el tamaño de la muestra que se necesita. Con un **muestreo aleatorio simple**, cualquier muestra posible de un tamaño particular tiene igual oportunidad de ser seleccionada. Por ejemplo, para seleccionar una muestra aleatoria simple de 100 estudiantes de todos los estudiantes en su escuela, podría asignar un número a cada uno y seleccionar la muestra sacando 100 de estos números de un sombrero. Siempre que el número de cada estudiante esté en el sombrero una sola vez, cada muestra de 100 estudiantes tiene igual oportunidad de ser seleccionada. Como una alternativa más rápida al uso del sombrero, podría seleccionar los números de los estudiantes con ayuda de una computadora o una calculadora que tenga integrado un *generador de números aleatorios*.

UN MOMENTO DE REFLEXIÓN

Busque la tecla de números aleatorios en una calculadora. (Casi todas las calculadoras científicas tienen una). ¿Qué sucede cuando la oprime? ¿Cómo podría utilizar la tecla de números aleatorios para seleccionar una muestra de 100 estudiantes?

Puesto que un muestro aleatorio simple proporciona a una muestra simple de un tamaño particular la misma oportunidad de ser elegida, es probable proporcionar una muestra representativa, siempre y cuando sea suficientemente grande. Entre mayor sea la muestra aleatoria simple, más probable es que sea representativa de la población.

EJEMPLO 3 Muestreo en un directorio telefónico

Usted podría llevar a cabo una encuesta de opinión en la cual la población sean todos los residentes en una ciudad. ¿Podría seleccionar una muestra aleatoria simple escogiendo nombres aleatoriamente del directorio telefónico local?

Solución Una muestra extraída de un directorio telefónico no es una muestra aleatoria simple de la población de la ciudad, ya que en los directorios invariablemente no hay muchos nombres y alguien que no aparezca no tiene oportunidad de ser elegido. Por ejemplo, el directorio telefónico no tendrá los nombres de dos o más personas que comparten el mismo número telefónico, sólo una de ellas está en la lista, o cuando una persona elige que su número telefónico no aparezca, o cuando dependen exclusivamente de un teléfono celular, o cuando una persona (tal como los vagabundos) no tienen teléfono.

Muestreo sistemático

El muestreo aleatorio simple es efectivo, pero en muchos casos podemos obtener resultados igualmente buenos con una técnica más sencilla. Suponga que está probando la calidad de los microcircuitos producidos por Intel. Conforme éstos salen de la línea de ensamblado, podría decidir probar cada 50 circuitos. Esto podría dar una muestra representativa ya que no existe razón para creer que cada 50 circuitos tenga una característica especial comparada con los otros circuitos. Este tipo de muestreo, en el que utilizamos un sistema, tal como la elección de cada 50 miembros de la población, se denomina **muestreo sistemático**.

EJEMPLO 4 Evaluación en un museo

Cuando el Museo Nacional del Aire y del Espacio necesitó probar ideas posibles para una nueva exhibición del sistema solar, un miembro del equipo entrevistó a una muestra de visitantes seleccionados por muestreo sistemático. Ella entrevistó a un visitante exactamente cada 15 minutos, seleccionando a quienquiera que fuese a entrar, en ese momento, a la exhibición actual del sistema solar. ¿Por qué cree que ella eligió el muestreo sistemático en lugar del muestreo aleatorio simple? ¿Es probable que el muestreo sistemático produzca una muestra representativa en este caso?

Solución El muestreo aleatorio simple, ocasionalmente, podría elegir dos visitantes que lleguen uno atrás de otro que el miembro del equipo no tendría tiempo de entrevistar a cada uno de ellos. El proceso sistemático de seleccionar un visitante cada 15 minutos, presentado en este ejemplo, previene este problema. Puesto que no existe razón para pensar que una persona que ingrese en un momento particular sea diferente de los que entran unos minutos antes o después, es probable que este proceso proporcione una muestra representativa de la población de visitantes durante el tiempo del muestreo.

EJEMPLO 5 Cuando el muestreo sistemático no funciona

Usted está realizando una encuesta de estudiantes en un dormitorio de una escuela mixta en el que los hombres tienen asignadas habitaciones con número impar y las mujeres habitaciones con número par. ¿Puede obtener una muestra representativa cuando elige cada 10 habitaciones?

Solución No. Si usted inicia con una habitación con número impar, cada 10 habitaciones también será una con número impar (tal como las habitaciones con número 3, 13, 23,...). De forma análoga, si inicia con una habitación con número par, cada 10 habitaciones también tendrá número par. Por tanto, obtendrá habitaciones con sólo hombres o con sólo mujeres, ninguna de ellas es representativa de la población de la escuela mixta.

UN MOMENTO DE REFLEXIÓN

Suponga que en el ejemplo 5 elige cada 5 habitaciones, en lugar de cada 10, ahora ¿la muestra sería representativa?

Apropósito...

El muestreo descrito en el ejemplo 4 fue emprendido antes de la construcción del Modelo a Escala del Sistema Solar Voyage, una exhibición permanente de un modelo a escala que se extiende a lo largo del paseo del Museo Nacional del Aire y del Espacio a la Torre Smithsonian. La fotografía siguiente muestra al hijo de uno de los creadores de la exhibición (también autor de este libro) tocando el modelo a escala del sol.



Muestreo de conveniencia

El muestreo sistemático es más sencillo que el muestreo aleatorio simple, pero también puede ser poco práctico en muchos casos. Por ejemplo, suponga que quiere conocer la proporción de estudiantes zurdos en su escuela. Sería un gran esfuerzo seleccionar una muestra aleatoria simple o una muestra sistemática, ya que ambas requieren elegir entre todos los estudiantes de la escuela. En contraste, sería más sencillo utilizar a los estudiantes en su clase de estadística —basta con pedir que levanten la mano los estudiantes que son zurdos—. Este tipo de muestra se denomina **muestra de conveniencia** ya que selecciona por conveniencia en lugar de hacerlo por un procedimiento más complejo. Al tratar de determinar la proporción de personas zurdas, el muestreo de conveniencia de su clase de estadística probablemente esté bien; no existe razón para pensar que la proporción de estudiantes sea diferente que la de la escuela. Pero si estuviese tratando de determinar la proporción de estudiantes con diferentes especialidades, esta muestra sería sesgada, ya que algunas especialidades requieren un curso de estadística y otras no. En general, el muestreo de conveniencia tiende a ser más propenso a ser sesgado que otras formas de muestreo.

EJEMPLO 6 Prueba del sabor de una salsa

Un supermercado necesita decidir si tener una nueva marca de salsa, por lo que ofrece una prueba gratis en un puesto en la tienda y pregunta a las personas si les gusta. ¿Qué tipo de muestreo se utiliza? ¿Es probable que la muestra sea representativa de la población de todos los clientes?

Solución La muestra de clientes que se paran a probar la salsa es un muestreo de conveniencia, ya que estas personas se encuentran en la tienda y son voluntarios para probar el producto nuevo. (Este tipo de muestreo de conveniencia, en el que las personas eligen o no ser parte de la muestra, también se denomina *muestreo de autoselección*. En la sección 1.4 estudiaremos más este tipo de muestreo). Tal muestra no es probable que sea representativa de la población de todos los clientes, ya que distintos tipos de personas podrían hacer sus compras en diferentes momentos (por ejemplo, los padres que están en casa es más probable que hagan las compras a mediodía que los padres que trabajan) y sólo las personas a quienes les gusta la salsa es probable que participen. Aun así, la información podría ser muy útil, ya que las opiniones de las personas que les gustan la salsa quizá sean más importantes en este caso.

Muestreo por conglomerado

El **muestreo por conglomerado** incluye la selección de *todos* los miembros de grupos (o *conglomerados*) seleccionados aleatoriamente. Imagine que trabaja para el Departamento de Agricultura y desea determinar el porcentaje de agricultores que utilizan técnicas de cultivo orgánicas. Sería difícil y costoso recolectar una muestra aleatoria simple o un muestreo sistemático ya que ambos requerirían la visita de muchas granjas que están ubicadas muy lejos unas de otras. Un muestreo de conveniencia de agricultores en un solo condado sería sesgado, ya que las técnicas de cultivo varían de región en región. Por tanto, podría decidir seleccionar de forma aleatoria unos cuantos condados en todo el país y encuestar a *todos* los agricultores de esos condados. Decimos que cada condado tiene un *conglomerado* de agricultores y la muestra consiste en *cada* agricultor que pertenece a los conglomerados seleccionados aleatoriamente.

EJEMPLO 7 Precios de gasolina

Usted quiere conocer el precio medio de la gasolina en las gasolineras localizadas en un radio de una milla de la ubicación de rentas de automóviles en los aeropuertos. Explique cómo podría utilizar el muestreo por conglomerados en este caso.

Solución Podría seleccionar aleatoriamente unos cuantos aeropuertos en todo el país. Para estos aeropuertos verificaría los precios de la gasolina en *cada* gasolinera en un radio de una milla de la ubicación de la renta de automóviles.

Apropósito...

Como lo manda la constitución de Estados Unidos, en realidad el voto para presidente lo realiza un pequeño grupo de personas denominadas *electores*. Cada estado puede seleccionar tantos electores como tenga miembros en el Congreso (contando senadores y representantes). Cuando usted emite un voto para presidente, en realidad emite un voto para el elector de su estado, cada uno de los cuales tiene el compromiso de votar por un candidato presidencial particular. Los electores emiten sus votos unas cuantas semanas después de la elección general.



Muestreo estratificado

Suponga que realiza una encuesta para pronosticar el resultado de la siguiente elección presidencial en Estados Unidos. La población bajo estudio son todos los votantes, por lo que podría seleccionar una muestra aleatoria simple de esta población. Sin embargo, ya que las elecciones presidenciales las decide el reparto de los votos electorales estado por estado, obtendría un mejor pronóstico si determina las preferencias de los votantes dentro de cada estado. Por tanto, su muestra global consiste de muestras aleatorias de cada uno de los 50 estados. En terminología estadística, las poblaciones de los 50 estados representan subgrupos, o **estratos**, de la población total. Puesto que su muestra total consiste en los miembros seleccionados de cada estrato, hemos utilizado **muestreo estratificado**.

EJEMPLO 8 Datos de desempleo

El Departamento del Trabajo de Estados Unidos encuesta cada mes 60 000 hogares para compilar su reporte de desempleo (vea ejemplo 2 en la sección 1.1). Para seleccionar estos hogares, este departamento primero agrupa las ciudades y condados en alrededor de 2 000 áreas geográficas. Luego, de forma aleatoria, selecciona hogares en cada una de estas áreas geográficas para encuestarlos. ¿Es un ejemplo de muestreo estratificado? ¿Cuáles son los estratos? ¿Por qué el muestreo estratificado es importante en este caso?

Solución La encuesta de desempleo es un ejemplo de muestreo estratificado ya que primero divide la población en subgrupos. Los subgrupos, o estratos, son las personas en las 2 000 regiones geográficas. En este caso el muestreo estratificado es importante, porque las tasas de desempleo probablemente son diferentes en las distintas regiones geográficas. Por ejemplo, las tasas de desempleo en la parte rural de Kansas podrían ser muy diferentes que en Silicon Valley. Al utilizar muestreo estratificado, el Departamento del Trabajo asegura que su muestra representa adecuadamente a todas las regiones geográficas.

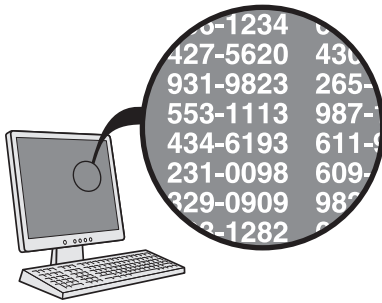
Resumen de métodos de muestreo

El recuadro siguiente y la figura 1.3 resumen los cinco métodos de muestreo que hemos analizado. Ningún método es “el mejor”, cada uno tiene sus propios usos. (Algunos estudios incluso combinan dos o más tipos de muestreo). Pero sin importar cómo se elija una muestra, tenga en mente las tres ideas clave siguientes:

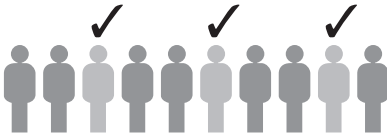
- Un estudio puede tener éxito sólo si la muestra es representativa de la población.
- Una muestra sesgada no es probable que sea una muestra representativa.
- Aunque una muestra sea bien elegida, puede resultar no representativa, por mala suerte, al sacar la muestra.

Métodos comunes de muestreo

- **Muestreo aleatorio simple:** elegimos una muestra de elementos de tal manera que cualquier muestra del mismo tamaño tenga igual oportunidad de ser seleccionada.
- **Muestreo sistemático:** utilizamos un sistema sencillo para seleccionar la muestra, tal como elegir a cada 10 o cada 50 elementos de la población.
- **Muestreo por conveniencia:** utilizamos una muestra que pueda ser conveniente de seleccionar.
- **Muestreo por conglomerados:** primero dividimos la población en grupos, o conglomerados, y seleccionamos al azar algunos de estos conglomerados. Luego obtenemos la muestra mediante la selección de *todos* los miembros en cada uno de los conglomerados seleccionados.
- **Muestreo estratificado:** utilizamos este método cuando nos interesan las diferencias entre subgrupos, o *estratos*, dentro de una población. Primero identificamos los estratos y luego sacamos una muestra aleatoria en cada uno de los estratos. La muestra total consiste en todas las muestras de cada uno de los estratos individuales.

**Muestreo aleatorio simple:**

Cada muestra del mismo tamaño tiene igual oportunidad de ser elegida. Con frecuencia las computadoras son utilizadas para generar números aleatorios.

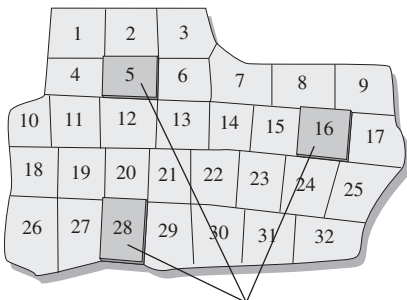
**Muestreo sistemático:**

Selecciona cada k elementos.

**Muestreo por conveniencia:**

Utiliza resultados que son fácilmente disponibles.

Elección de distritos en el condado de Carson



Entrevista a todos los votantes en los distritos sombreados.

Muestreo por conglomerados:

Divide la población en conglomerados, selecciona algunos de esos conglomerados y luego toma a todos los elementos de los conglomerados seleccionados.

**Muestreo estratificado:**

Divide a la población en al menos dos estratos, luego se saca una muestra de cada uno de ellos.

Figura 1.3

EJEMPLO 9 Métodos de muestreo

Identifique el tipo de muestreo utilizado en cada uno de los casos siguientes:

- La cosecha de manzanas en un huerto se recolecta en 1 200 canastos. Un inspector de agricultura selecciona 25 canastos al azar y luego revisa si cada manzana en estos canastos tiene gusanos.
- Una investigadora educativa quiere conocer si, en una escuela en particular, los hombres y las mujeres tienden a realizar más preguntas en clase. De los 10 000 estudiantes en el colegio, ella entrevista a 50 hombres y a 50 mujeres elegidas al azar.
- Al tratar de aprender si los sistemas planetarios son comunes, los astrónomos realizan un sondeo buscando planetas entre 100 de las estrellas más cercanas.
- Para determinar quién ganará balones de fútbol autografiados, un programa de cómputo selecciona aleatoriamente los números de boleto de 11 personas en un estadio lleno con gente.

Solución

- La inspección de manzanas es un ejemplo de muestreo por conglomerados, ya que el inspector inicia con un conjunto de conglomerados (los canastos) elegidos de manera aleatoria y luego revisa todas las manzanas en los conglomerados seleccionados.
- Para este estudio los grupos de hombres y de mujeres representan dos estratos diferentes, por lo que es un ejemplo de muestreo estratificado.
- Los astrónomos centran su atención en estrellas cercanas, ya que son más sencillas de estudiar, por lo que es un ejemplo de muestreo por conveniencia.
- Puesto que la computadora selecciona al azar los 11 boletos, cada boleto tiene igual oportunidad de ser elegido. Es un ejemplo de muestreo aleatorio simple.

Sección 1.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Censo y muestra.** ¿Qué es un censo, qué es una muestra y cuál es la diferencia entre ellos?
- Muestra sesgada.** En una encuesta antes de elecciones, se saca una muestra aleatoria de una lista de demócratas registrados. ¿Existe algo erróneo en este método de muestreo?
- Muestreo por conglomerados y estratificado.** Los muestreos por conglomerados y estratificado incluyen la selección de sujetos en subgrupos de la población. ¿Cuál es la diferencia entre estos dos tipos de muestreos?
- Muestreo de estudiantes.** Uno de los autores realizó un estudio registrando si cada estudiante en su clase era diestro. El objetivo era sacar una conclusión acerca de la proporción de estudiantes de la universidad que eran diestros. ¿Qué tipo de muestra obtuvo? ¿Quizás esta muestra fue sesgada? ¿Por qué sí o por qué no?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Edad de graduación.** Para un proyecto de la clase de estadística, yo realicé un censo para determinar la edad media de estudiantes cuando ellos obtienen su grado de licenciatura.
- Muestreo de conveniencia.** Para un proyecto de la clase de estadística, utilicé un muestreo de conveniencia, pero los resultados podrían no ser significativos.
- Muestra sesgada.** El estudio puede estar sesgado, ya que concluyó que 75% de los estadounidenses tienen más de 6 pies de altura.
- Condenados a muerte.** En la actualidad existen 3 366 convictos con pena de muerte (con base en datos de 2007 de la Oficina de Estadísticas de Justicia). Obtuvimos una muestra aleatoria simple de esos reos compilando una lista numerada, luego utilizamos una computadora para generar aleatoriamente 20 números entre 1 y 3 366 y luego seleccionamos a los reos que correspondían a los números generados.

Conceptos y aplicaciones

Censo. En los ejercicios 9 a 12 determine si un censo es práctico en la situación que se describe. Explique su razonamiento.

- Altura de jugadores.** Usted quiere determinar la altura media de todos los jugadores de básquetbol del equipo de los Lakers de Los Ángeles.

10. **Estatura en la preparatoria.** Quiere determinar la estatura media de todos los jugadores de básquetbol en las preparatorias de Estados Unidos.
11. **Puntuación de IQ.** Quiere determinar la puntuación media de IQ de todos los profesores de estadística en Estados Unidos.
12. **Edades de los profesores.** Quiere determinar la edad media de todos los profesores de estadística en la Universidad de Colorado.

¿Muestras representativas? En los ejercicios del 13 al 16 identifique la muestra, la población y el método de muestreo. Luego comente si es probable que la muestra sea representativa de la población.

13. **Duración en el Senado.** Un científico de la política selecciona aleatoriamente 4 de los 100 senadores que actualmente están en el Congreso y luego determina el tiempo que han estado en servicio.
14. **Súper Tazón.** Durante el juego del Súper Tazón, Nielsen Media Research lleva a cabo una encuesta en 5 108 hogares seleccionados aleatoriamente y determina que 44% de ellos tienen la televisión sintonizada en el Súper Tazón.
15. **Clonación.** En una encuesta de Gallup a 1 012 adultos estadounidenses elegidos aleatoriamente, 89% dijo que la clonación de humanos no debe permitirse.
16. **Encuesta por correo.** Una estudiante graduada en la Universidad de Newport realizó un proyecto de investigación acerca de cómo los adultos estadounidenses se comunican. Ella inicia con una encuesta enviada por correo a 500 de los adultos que conoce. Les pide que regresen por correo una respuesta a esta pregunta: “¿Prefiere utilizar correo electrónico o el correo “caracol” (el servicio postal de Estados Unidos)?”. Obtiene 65 respuestas, con 42 de ellas indicando que prefieren el correo caracol.

Evaluación de elecciones de muestreo. Los ejercicios 17 y 18 describen el objetivo de un estudio y luego le ofrecen cuatro posibles muestras. En cada caso decida cuál muestra es más probable que sea una muestra representativa y explique por qué. Después explique por qué cada una de las otras opciones es probable que *no* sea una muestra representativa para el estudio.

17. **Deuda en tarjeta de crédito.** Usted quiere determinar el monto promedio (medio) de tarjetas de crédito de adultos consumidores en Florida:
 - a. Los conductores de Florida que poseen y tienen registrados vehículos Land Rover.
 - b. Los primeros 1 000 residentes de Florida listados en el directorio de Fort Lauderdale.
 - c. Los primeros 1 000 residentes de Florida en una lista completa de todos los residentes de números telefónicos de Florida.
 - d. Los residentes de Florida que regresen por correo una encuesta impresa en el periódico *Miami Herald*.

18. **Votantes en California.** Usted quiere llevar a cabo una encuesta para determinar la proporción de votantes elegibles en California que probablemente voten por el candidato demócrata a la presidencia en la siguiente elección.

- a. Todos los votantes elegibles en el condado de San Diego.
- b. Todos los votantes elegibles en la ciudad de Sonoma.
- c. Todos los votantes elegibles que respondan a una encuesta por internet de AOL (America OnLine).
- d. Cada 1 000 personas en una lista completa de todos los votantes elegibles en California.

Sesgo. En las situaciones descritas en los ejercicios del 19 al 22, ¿existe sesgo?

19. **Crítica de película.** Una crítica de cine para ABC News le da su opinión de la última película de Disney, que también pasó por ABC.
20. **Revisión de automóviles.** La revista *Consumer Reports*, que no acepta productos o publicidad gratuita de nadie, imprime un estudio de automóviles nuevos.
21. **Soya genéticamente modificada.** Monsanto contrata científicos universitarios independientes para determinar si su nueva soya modificada genéticamente plantea algún peligro para el medio ambiente.
22. **Fondos para estudio de medicamentos.** El *Journal of the American Medical Association* publica un artículo que evalúa un medicamento y algunos de los médicos que escriben el artículo recibieron fondos de la compañía farmacéutica que lo produce.

Métodos de muestreo. En los ejercicios del 23 al 38 identifique cada muestra como aleatoria simple, sistemática, de conveniencia, estratificada o de conglomerado. En cada caso establezca si considera que el método probablemente dará una muestra representativa o una sesgada y explique por qué.

23. **Ensayo clínico.** En la fase II de prueba de un medicamento nuevo diseñado para aumentar el conteo de glóbulos rojos en la sangre, una investigadora consigue sobres con los nombres y direcciones de todos los sujetos tratados. Ella quiere aumentar la dosis en una submuestra de 12 sujetos, por lo que mezcla todos los sobres en una bandeja y luego saca 12 de esos sobres para identificar a los sujetos a quienes les aumentará la dosis.
24. **Punto de revisión de sobriedad.** Uno de los autores fue observado en un punto policial de verificación de sobriedad, en el que cada quinto conductor era detenido y revisado.
25. **Encuestas de salida.** En el día de la elección presidencial los noticiarios organizaron una encuesta de salida en la que casillas de votación específicas fueron seleccionadas aleatoriamente y todos los votantes eran encuestados conforme salían de las instalaciones.

- 26. Educación y deportes.** Un investigador para la compañía de artículos deportivos Spalding estudia la relación entre el nivel educativo y la participación en algún deporte. Él lleva a cabo una encuesta de 40 golfistas elegidos al azar, 40 tenistas elegidos al azar y 40 nadadores elegidos al azar.
- 27. Ergonomía.** Un estudiante de ingeniería mide la fuerza de los dedos utilizados para oprimir botones haciendo la prueba con los miembros de la familia.
- 28. Evasión de impuestos.** Un investigador del Servicio de Ingresos Internos investiga la evasión en los reportes de impuestos a los ingresos encuestando a todos los camareros y las camareras en 20 restaurantes elegidos aleatoriamente.
- 29. Encuesta en MTV.** Un experto en marketing para MTV planea una encuesta en la que 500 personas se seleccionarán aleatoriamente de cada uno de los grupos de edad 10-19, 20-29, y así sucesivamente.
- 30. Información de tarjetas de crédito.** Uno de los autores encuestó a todos sus estudiantes para obtener datos muestrales del número de estudiantes que tenían tarjeta de crédito.
- 31. Recaudación de fondos.** Los recaudadores de fondos para el Colegio de Newport prueban una nueva campaña de telemarketing obteniendo una lista alfabética de todos los alumnos y llaman a cada 100 en esa lista.
- 32. Encuesta telefónica.** En una encuesta de Gallup a 1 059 adultos, los sujetos entrevistados fueron seleccionados usando una computadora para elegir números telefónicos generados al azar.
- 33. Investigación de mercado.** Una investigadora de mercado dividió a todos los residentes de California en categorías: desempleados, empleados de tiempo completo y empleados de medio tiempo. Ella está encuestando a 50 personas de cada categoría.
- 34. Estudiantes bebedores.** Motivado por la muerte de un estudiante en una borrachera, el Colegio de Newport realizó un estudio de los estudiantes que beben, seleccionó al azar 10 diferentes grupos y entrevistó a todos los estudiantes de esos grupos.
- 35. Encuesta en una revista.** La revista *People* eligió a sus “famosos mejor vestidos” compilando las respuestas que sus lectores enviaron por correo a una encuesta impresa en la revista.
- 36. Trasplantes de corazón.** Un investigador médico en la Universidad Johns Hopkins obtiene una lista numerada de todos los pacientes que esperan por un trasplante de corazón, utiliza una computadora para generar 50 números aleatorios y luego selecciona los pacientes que correspondan a los 50 números.

37. Control de calidad. Una muestra de CD producidos se obtiene por medio de una computadora que genera aleatoriamente un número entre 1 y 1 000 para cada CD y luego se selecciona el CD si el número generado es 1 000.

38. Cinturones de seguridad. Cada 500 cinturones de seguridad se prueban mediante esfuerzo hasta que fallen.

Selección de un método de muestreo. Para cada uno de los ejercicios del 39 al 42, sugiera un método de muestreo que produzca una muestra representativa. Explique por qué seleccionó este método sobre otros métodos.

39. Elección estudiantil. Quiere pronosticar al ganador en la próxima elección para presidente estudiantil.

40. Tipo de sangre. Quiere determinar el porcentaje de personas en este país en cada uno de los cuatro grupos sanguíneos principales (A, B, AB y O).

41. Muertes por cardiopatías. Quiere determinar el porcentaje de muertes anuales debidas a cardiopatías.

42. Mercurio en atunes. Quiere determinar el promedio de mercurio contenido en el atún que consumen los residentes de Estados Unidos.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 1 en www.aw.com/bbt.

43. Encuesta de opinión pública. Utilice la información disponible en el sitio web de una compañía de encuestas, tal como Gallup, Harris, Pew o Yankelovich, para responder las preguntas siguientes:

- a. Exactamente, ¿cómo se selecciona una muestra de individuos?
- b. Con base en lo que ha aprendido, ¿considera que los resultados de la encuesta son confiables? Si es así, ¿por qué? Si no, ¿por qué no?

44. Muestra de desempleo. Utilice la página web de la Oficina de Estadísticas de Trabajo para encontrar los detalles de cómo esta oficina selecciona la muestra de hogares en su encuesta mensual. Escriba un breve resumen del procedimiento y por qué es probable que arroje una muestra representativa.

45. Votación selectiva. Los premios de la Academia, el Trofeo Heisman y la lista de éxitos de venta de *New York Times* son sólo tres ejemplos de casos en los que las selecciones están determinadas por votos de individuos especialmente elegidos. Vea uno de estos procesos de selección y describa quiénes votan y cómo son elegidas estas personas. Analice las fuentes de sesgo en el proceso.

EN LAS NOTICIAS

- 46. Muestreo en los noticiarios.** Encuentre un reportaje de una noticia reciente acerca de una investigación estadística que sea de su interés. Escriba un breve resumen de cómo fue elegida la muestra para el estudio y analice brevemente si cree que la muestra fue representativa de la población bajo estudio.
- 47. Muestra en encuesta de opinión.** Determine un reportaje de una noticia reciente de una encuesta de opinión llevada a cabo por una compañía de noticias (tal como Gallup, Harris, *USA Today*, *New York Times* o CNN). Describa brevemente la muestra y cómo se eligió. ¿La muestra se seleccionó de manera tal que sea probable que se le introduzca un sesgo? Explique.

- 48. Encuestas políticas.** Encuentre resultados de una encuesta reciente llevada a cabo por una organización política (tal como el partido Republicano o el Demócrata o una organización que busque influir al Congreso sobre un tema en particular). Describa brevemente la muestra y cómo se seleccionó. ¿La muestra fue elegida de tal manera que es probable introducir sesgo? ¿Debería estar más preocupado acerca del sesgo en tal encuesta de lo que estaría en una encuesta realizada por una organización de noticias? Explique.

1.3 Tipos de estudios estadísticos

Los estudios estadísticos se realizan de muchas formas distintas. En todos los casos las personas, los animales (u otros seres vivos) u objetos seleccionados para la muestra se denominan **sujetos** del estudio. Si los sujetos son personas, es común referirse a ellos como los **participantes** en el estudio.

Definición

Los **sujetos** de un estudio son las personas, los animales (u otros seres vivos) u objetos seleccionados para la muestra. Si los sujetos son personas, también se les puede llamar **participantes** en el estudio.

Existen dos tipos básicos de estudios estadísticos: estudios de observación y experimentos. En un **estudio de observación** observamos o medimos características específicas tratando de ser cuidadosos en evitar la influencia o modificación de las características en observación. Los índices de audiencia Nielsen son un ejemplo de un estudio de observación, ya que Nielsen utiliza sus “medidores de personas” para *observar* lo que están viendo en el televisor los participantes, pero no trata de influir en lo que ven.

Note que un estudio de observación puede incluir actividades que van más allá de la definición usual de *observación*. La medición de pesos de personas requiere interactuar con ellas, al igual que al preguntarles para colocarlas en alguna escala. Pero en estadística consideramos estas medidas como observaciones, ya que las interacciones no cambian el peso de las personas. De forma análoga, una encuesta de opinión en la que los investigadores realizan entrevistas a profundidad es considerada de observación, siempre que los investigadores sólo intenten conocer las opiniones de las personas y no cambiarlas.

En contraste, considere un estudio médico diseñado para probar si dosis mayores de vitamina C ayudan a prevenir resfriados. Para llevar a cabo el estudio, los investigadores deben pedir a algunas personas en la muestra que diariamente tomen dosis mayores de vitamina C. Este tipo de investigación estadística se denomina **experimento**. El propósito de un experimento es estudiar los efectos de algún **tratamiento**, en este caso una dosis diaria mayor de vitamina C.

Puedes observar mucho tan sólo mirando.

—Yogi Berra

Dos tipos básicos de investigación estadística

Existen dos tipos básicos de investigación estadística:

1. En un **estudio de observación** los investigadores observan o miden características de los sujetos, pero no intentar influir o modificar estas características.
2. En un **experimento** los investigadores aplican algún **tratamiento** y observan sus efectos en los participantes en el experimento.

EJEMPLO 1 Tipo de estudio

Identifique si se trata de un estudio de observación o un experimento.

- a. El estudio de la vacuna contra la polio de Salk (vea el estudio de caso en la página 8).
- b. Una encuesta en la que se pregunta a la gente por quién planean votar en la siguiente elección.

Solución

- a. El estudio de la vacuna contra la polio de Salk fue un *experimento*, ya que los investigadores probaron un tratamiento —en este caso, la vacuna— para ver si reducía la incidencia de la polio.
- b. La encuesta es un *estudio de observación* ya que intenta determinar la preferencia de voto pero no intenta influir en los votos.

Identificación de las variables

Por lo común los estudios estadísticos, ya sean de observaciones o de experimentos, intentan medir lo que denominamos **variables de interés**. El término *variable* describe un elemento o cantidad que puede variar o tomar valores diferentes, y las variables de interés son aquellas sobre las que buscamos aprender. Por ejemplo, las variables de interés en los estudios de hábitos de televidentes de Nielsen incluyen *programa que están viendo* y *número de espectadores*. La variable *programa que están viendo* puede tomar valores diferentes tal como *Súper Tazón*, *60 minutos* o *Lost*. La variable *número de espectadores* depende de la popularidad de un programa. En esencia, los datos de cualquier investigación estadística son los valores diferentes de las variables de interés.

En el caso que considere causa y efecto, subdivida las variables de interés en dos categorías. Por ejemplo, cada persona en el estudio de la vitamina C y resfriados puede tomar una dosis diferente de vitamina C y podría terminar con un número diferente de resfriados en algún periodo. Puesto que trata de saber si la vitamina C causa una disminución en el número de resfriados, decimos que la *dosis diaria de vitamina C* es una **variable explicatoria**, puede explicar o causar un cambio en el número de resfriados. De forma análoga, decimos que el *número de resfriados* es una **variable de respuesta**, ya que esperamos que responda a cambios en la variable explicatoria (la dosis de vitamina C).

Definiciones

Una **variable** es cualquier elemento o cantidad que puede variar en valores diferentes.

Las **variables de interés** en una investigación estadística son los elementos o cantidades que el estudio busca medir.

Cuando busca causa y efecto, una **variable explicatoria** es una variable que puede explicar o causar el efecto, mientras que una **variable de respuesta** responde a cambios en la variable explicatoria.

EJEMPLO 2 Identificar las variables

Identifique las variables de interés en cada estudio:

- a. El estudio de la vacuna contra la polio de Salk.
- b. Una encuesta en la que a la gente le preguntaron por quién planea votar en la elección siguiente.

Solución

- a. Las dos variables de interés en el estudio de la vacuna de Salk son la *vacuna* y la *polio*. Son variables ya que pueden tomar dos valores diferentes: a un niño se le da o no se le da la vacuna y él contrae o no contrae la polio. En este caso, puesto que el estudio busca determinar si la vacuna previene la polio, decimos que la *vacuna* es la variable explicatoria (puede explicar un cambio en la incidencia de la polio) y la *polio* es la variable de respuesta (se supone que cambia en respuesta a la vacuna).
- b. Las variables de interés en la encuesta previa a la elección podrían llamarse *nombre del candidato*, ya que la gente debe elegir uno de los candidatos por el que planea votar y *proporción de partidarios*, que es la fracción de la gente encuestada que planea votar por un candidato particular. (En este estudio no están involucrados causa ni efecto, por lo que no necesitamos decidir si la variable es explicatoria o de respuesta).

Estudios de observación

Los estudios de observación que hemos analizado hasta aquí, tal como los índices de audiencia de Nielsen, las encuestas de opinión y la determinación de las estaturas de los estudiantes, son estudios en los que, por lo común, todos los datos se recolectan casi al mismo tiempo. Sin embargo, algunas veces los estudios de observación examinan datos pasados o están diseñados para examinar datos en un periodo largo.

Un **estudio en retrospectiva** (en ocasiones denominado *estudio con control de caso*) es un estudio de observación que utiliza datos del pasado, tal como registros oficiales o entrevistas anteriores, para aprender acerca de un tema de interés. Los estudios en retrospectiva son especialmente valiosos cuando puede ser poco práctico o no ético realizar un experimento. Por ejemplo, suponga que quiere aprender cómo el alcohol consumido durante el embarazo afecta a los bebés recién nacidos. Puesto que ya sabe que el consumo de alcohol durante el embarazo puede ser dañino, no sería ético pedir a las madres embarazadas probar el “tratamiento” de consumo de alcohol. Sin embargo, puesto que muchas madres consumen alcohol en embarazos previos (ya sea antes de que fuesen conocidos los peligros o porque deciden ignorar los peligros), podemos realizar un estudio en retrospectiva en el que comparamos hijos nacidos de esas madres con niños nacidos de madres que no consumieron alcohol.

En ocasiones, la información que necesitamos para llegar a conclusiones claras no está disponible en los registros pasados. En esos casos, los investigadores pueden preparar un **estudio en prospectiva** (estudio *longitudinal*) diseñado para recolectar observaciones en el futuro de grupos que comparten factores comunes. Un ejemplo clásico de un estudio en prospectiva es el Estudio de Salud de las Enfermeras de Harvard, el cual se inició en 1976 para recolectar datos acerca de cómo estilos de vida diferente afectan la salud de las mujeres (vea la sección “Hablemos de salud pública” al final de este capítulo). El estudio, que aún continúa en la actualidad, ha seguido a miles de enfermeras durante más de tres décadas, recolectando información acerca de sus estilos de vida y su salud.

Variaciones en estudios de observación

En los estudios de observación más comunes los datos son recolectados de una vez (o tan cercano a esto como sea posible). Dos variaciones en estudios de observación también son comunes:

1. Un **estudio en retrospectiva** (o **control de caso**) utiliza la información del pasado, tal como registros oficiales o entrevistas pasadas.
2. Un **estudio en prospectiva** (o **longitudinal**) tiene un diseño para recolectar datos en el futuro de grupos que comparten factores comunes.

EJEMPLO 3 Estudio de observación

Usted quiere saber si a los niños nacidos prematuramente les va tan bien en la escuela elemental como a los niños nacidos en término. ¿Qué tipo de estudio debe hacer?

Solución En este caso la única opción real es hacer un estudio de observación en retrospectiva. Recolectaría los datos de nacimientos anteriores y compararía el desempeño en la escuela elemental de los nacidos prematuramente con los nacidos en término.

Experimentos

Puesto que los experimentos requieren de intervención activa, tal como la aplicación de un tratamiento, debemos tener cuidado especial en asegurar que están diseñados de manera que proveerán de la información que buscamos. Revisemos algunos de los temas que surgen en el diseño de experimentos.

Necesidad de controles

Considere un experimento que suministra vitamina C a algunos participantes en un estudio para determinar su efecto sobre los resfriados. Suponga que la gente que diario toma vitamina C tiene un promedio de 1.5 resfriados en un periodo de tres meses. ¿Cómo los investigadores pueden saber si los participantes hubiesen tenido más resfriados sin la vitamina C? Para responder este tipo de pregunta, los investigadores deben llevar a cabo sus experimentos con dos (o más) grupos de sujetos: un grupo que diario toma grandes dosis de vitamina C y otro grupo que no lo hace. Como estudiaremos dentro de poco, en la mayoría de los casos es importante que los participantes sean asignados de manera aleatoria a cada uno de los dos grupos.

El grupo de personas que se asignan aleatoriamente para tomar vitamina C se denomina **grupo de tratamiento**, ya que sus miembros reciben el tratamiento que será probado (vitamina C). El grupo de personas que *no* toma la vitamina C, se denomina **grupo de control**. Los investigadores pueden tener confianza que la vitamina C es un tratamiento efectivo sólo si las personas en el grupo de tratamiento tienen un número significativamente menor de resfriados que la gente en el grupo de control. El grupo de control recibe su nombre porque ayuda a controlar la forma de interpretar los resultados experimentales.

Grupos de tratamiento y de control

El **grupo de tratamiento** en un experimento es el grupo de sujetos quienes reciben el tratamiento que será probado.

El **grupo de control** en un experimento es el grupo de sujetos que *no* reciben el tratamiento que será probado.

En la mayoría de los casos es importante elegir a los miembros de los dos grupos mediante una selección aleatoria del conjunto total de sujetos disponibles.

EJEMPLO 4 Tratamiento y control

Revise el estudio de caso sobre la vacuna contra la poliomielitis de Salk (página 8). ¿Cuál fue el tratamiento? ¿Qué grupo de niños constituyó el grupo de tratamiento? ¿Cuál grupo constituyó el grupo de control?

Solución El tratamiento fue la vacuna de Salk. El grupo de tratamiento consistió en los niños que recibieron la vacuna de Salk. El grupo de control consistió en los niños a quienes no se les aplicó la vacuna de Salk y en su lugar se les aplicó una inyección de agua con sal.

EJEMPLO 5 Tratamiento Mozart

Un estudio divide a los estudiantes de una universidad en dos grupos. Un grupo escuchó a Mozart u otra música clásica, antes de que se le asignara una tarea específica, y al otro grupo

sólo se le asignó la tarea. Los investigadores encontraron que quienes escucharon la música clásica realizaron la tarea un poco mejor, pero sólo si ellos hicieron la tarea unos cuantos minutos después de haber escuchado la música (los dos grupos realizaron las tareas dadas después). Identifique el tratamiento y los grupos de tratamiento y de control.

Solución El tratamiento fue la música clásica. El grupo de tratamiento consistió en los estudiantes, quienes escucharon la música. El grupo de control consistió en los estudiantes que no escucharon la música.

Variable de confusión

Utilizar grupos de control ayuda a asegurar que contamos con variables conocidas que podrían afectar los resultados de un estudio. Sin embargo, los investigadores pueden no ser conscientes o ser incapaces de contar otras variables importantes. Considere un experimento en el que un maestro de estadística busca determinar si quienes estudian en colaboración (en grupos de estudio con otros estudiantes) obtienen calificaciones más altas que quienes lo hacen de manera aislada. El maestro selecciona a cinco jóvenes que estudiarán en colaboración (el grupo de tratamiento) y otros cinco que estudiarán de manera aislada (grupo de control). Para asegurar que los estudiantes tienen capacidades similares y estudiarán con diligencia, el maestro selecciona sólo estudiantes con altos promedios de calificación. Al final del semestre el maestro encuentra que quienes estudiaron en colaboración obtuvieron calificaciones más altas.

Las variables de interés para este estudio son *el estudio en colaboración* (lo hagan así o no) y la *calificación final*. Pero suponga que, sin conocimiento del maestro, todos los estudiantes del grupo en colaboración viven en un dormitorio donde una restricción asegura que ellos duermen lo suficiente. Este hecho introduce una nueva variable, que podría llamarse *cantidad de sueño*, la cual podría explicar parcialmente los resultados. En otras palabras, la conclusión del experimento *parece* apoyar los beneficios del estudio en colaboración, pero esta conclusión no está justificada porque el maestro no tomó en cuenta cuánto durmieron los estudiantes.

En términos estadísticos dicho estudio sufre de **confusión**. Las calificaciones mayores podrían deberse a la variable de interés (*estudio en colaboración*) o bien a los diferentes horarios para dormir o bien a una combinación de ambos. Puesto que el maestro no tomó en cuenta las diferencias en las horas que se durmió, la *cantidad de sueño* es una **variable de confusión** para este estudio. Posiblemente usted piense en otras variables de confusión potenciales que pueden afectar un estudio como éste.

A propósito. .

El llamado efecto Mozart asegura que escuchar a Mozart puede hacer niños más inteligentes. El supuesto efecto generó toda una industria de productos Mozart para niños. El estado de Georgia ha empezado a repartir CD de Mozart a las nuevas mamás. Sin embargo, estudios más recientes del efecto Mozart no han sido capaces de probar tal efecto.



Definición

Un estudio sufre de **confusión** si los efectos de variables diferentes se mezclan de modo que no puede determinarse los efectos específicos de las variables de interés. Las variables que llevan a la confusión se denominan **variables de confusión**.

ESTUDIO DE CASO

Confusión en resultados con medicamentos

La compañía Pfizer desarrolló un medicamento nuevo (denominado fluconazole) diseñado para prevenir micosis en pacientes de hospital. Varios estudios determinaron que el medicamento nuevo era más efectivo que un medicamento anterior (denominado amphotericin B). Sin embargo, un análisis posterior, realizado por otros investigadores, determinó que el medicamento anterior había sido administrado oralmente, cuando se suponía que debía haberse administrado por medio de una inyección. Esto introdujo una fuente de confusión en los estudios: los investigadores originales pensaron que los resultados mostraron que el medicamento nuevo era más efectivo que el anterior, pero no habían tomado en cuenta los efectos de confusión de cómo había sido administrado el medicamento anterior. Una vez que los nuevos investigadores tomaron en cuenta este efecto (comparando sólo los casos en los que ambos medicamentos se administraron de manera correcta), determinaron que el medicamento nuevo no era más efectivo que el medicamento anterior.

Asignación de grupos de tratamiento y de control

Como ilustra el experimento del estudio en colaboración, los resultados casi siempre sufren de confusión, si los grupos de tratamiento y de control difieren en formas importantes (distintas a que reciban o no el tratamiento). Por lo común, los investigadores emplean dos estrategias para prevenir tales diferencias y de este modo asegurar que los grupos de tratamiento y de control puedan compararse sin problemas. Primero, asignan integrantes de los grupos de tratamiento y de control *al azar*, lo cual quiere decir que utilizan una técnica diseñada para asegurar que cada participante tiene igual oportunidad de ser asignado a cualquiera de los grupos. Cuando los participantes se asignan de manera aleatoria, es menos probable que las personas en los grupos de tratamiento y de control difieran de alguna manera que afecte los resultados del estudio.

Segundo, los investigadores tratan de asegurar que los grupos de tratamiento y de control sean suficientemente grandes. Por ejemplo, en el experimento del estudio en colaboración, la inclusión de 50 estudiantes en cada grupo, en lugar de 5, habría hecho menos probable que todos los estudiantes en un grupo vivieran en un dormitorio especial.

Estrategias para la selección de los grupos de tratamiento y de control

1. **Seleccionar los grupos al azar.** Asegurar que los sujetos del experimento sean asignados al grupo de tratamiento o de control al azar significa que cada sujeto tiene igual oportunidad de ser asignado a cualquier grupo.
2. **Usar grupos suficientemente grandes.** Asegurar que tanto el grupo de tratamiento como el de control sean suficientemente grandes para reducir la probabilidad de que difieran de una manera significativa (salvo por el hecho que a un grupo se le da el tratamiento y al otro no).

EJEMPLO 6 Grupos de estudio Salk

Brevemente explique cómo se utilizaron las dos estrategias para la selección de los grupos de tratamiento y de control para el estudio de la vacuna contra la polio de Salk.

Solución Participaron un total de alrededor de 400 000 niños en el estudio, a la mitad se le aplicó una inyección de la vacuna de Salk (el grupo de tratamiento) y la otra mitad recibió una inyección de agua salada (el grupo de control). La primera estrategia se implementó seleccionando de manera aleatoria, de entre todos los niños, a los niños para los dos grupos. La segunda estrategia se aplicó mediante el uso de un gran número de participantes (200 000 en cada grupo) de modo que es poco probable que por accidente los dos grupos difieran.

El efecto placebo

Cuando un experimento involucra a gente, los efectos pueden ocurrir porque la gente se sabe parte del experimento. Por ejemplo, suponga que usted está probando la efectividad de un nuevo medicamento antidepresivo. Usted encuentra 500 personas que sufren de depresión y de manera aleatoria las divide en un grupo de tratamiento que recibe el nuevo medicamento y un grupo de control que no lo recibe. Unas semanas después, entrevistas con los pacientes muestran que la gente en el grupo de tratamiento tiende a sentirse mucho mejor que la gente en el grupo de control. ¿Puede concluir que el nuevo medicamento funciona?

Por desgracia, es muy posible que el humor de la persona que recibe el medicamento mejore por el simple hecho de que fue agradable recibir alguna clase de tratamiento, lo cual significa que usted no puede estar seguro de que el medicamento realmente sea útil. Este tipo de efecto, en el que las personas mejoran porque creen que han recibido un tratamiento útil, es el **efecto placebo**. (La palabra *placebo* proviene del latín *placere*).

Para distinguir entre los resultados provocados por el efecto placebo y los que en realidad son debidos al tratamiento, los investigadores tratan de asegurarse que los participantes no saben si ellos son parte del grupo de tratamiento o del de control. Para hacerlo dan a la gente en el grupo de control un **placebo**: algo que parezca o se sienta igual al tratamiento que se prueba,

*Una gran imaginación provocó el [efecto placebo]. . .
Todo el mundo siente su impacto, pero algunos son derrumbados por él. . .
[los doctores] saben que existen hombres para los cuales la simple vista de la medicina es curativa.*

—Michel de Montaigne
(1533-1592), filósofo francés

pero que carezca de sus ingredientes activos. Por ejemplo, en una prueba de un medicamento que viene en forma de píldoras, el placebo podría ser una píldora de la misma forma y tamaño que contenga azúcar en lugar del medicamento real. En una prueba de una vacuna inyectada, el placebo podría ser una inyección que sólo contenga una solución salina (agua con sal) en lugar de la vacuna real. En una prueba reciente de la efectividad de la acupuntura, el placebo consistió en un tratamiento con agujas como en la acupuntura real, salvo que las agujas no se colocaban en los lugares especiales que los acupunturistas aseguran son importantes.

Mientras los participantes no sepan si recibieron el tratamiento real o un placebo, el efecto placebo debería afectar de la misma forma a los grupos de tratamiento y de control. Si los resultados de los dos grupos son significativamente diferentes, las diferencias pueden ser atribuidas al tratamiento. Por ejemplo, en el estudio del medicamento antidepresivo, concluiríamos que el medicamento fue efectivo sólo si el grupo de control recibió un placebo y los miembros del grupo de tratamiento mejoraron mucho más que los miembros del grupo de control. Para un mejor control, algunos experimentos utilizan tres grupos, un grupo de tratamiento, un grupo placebo y un grupo de control. Al grupo placebo se le da un placebo, mientras que al grupo de control no se le da nada.

Definiciones

Un **placebo** carece de ingredientes activos de un tratamiento que se prueba en un estudio, pero parece o se siente como el tratamiento, de modo que los participantes no pueden distinguir si recibieron el placebo o el tratamiento real.

El **efecto placebo** es la situación en que los pacientes mejoran sólo porque creen que recibieron un tratamiento útil.

Observación: aunque los participantes no deben saber si pertenecen al grupo de tratamiento o al grupo de control, por razones éticas es muy importante que se les diga que a algunos de ellos se les suministrará un placebo, mientras que otros recibirán el tratamiento real.

UN MOMENTO DE REFLEXIÓN

En las décadas pasadas, con frecuencia, los investigadores les decían a todos los participantes en un estudio que recibirían el tratamiento real, pero en realidad se daba un placebo a la mitad de ellos. Analice algunas razones de por qué ahora esto se consideraría poco ético. ¿Debe permitirse a los investigadores utilizar resultados de estudios pasados que no cumplen con los criterios éticos actuales? Defienda su opinión.

EJEMPLO 7 Vacuna placebo

¿Cuál fue el placebo en el estudio de la vacuna contra la polio de Salk? ¿Por qué los investigadores utilizaron un placebo en este experimento?

Solución El placebo fue la inyección de agua con sal suministrada a los niños en el grupo de control. Para entender por qué los investigadores utilizaron un placebo para el grupo de control, suponga que *no* se hubiese utilizado un placebo. Cuando las mejoras fuesen observadas en el grupo de tratamiento, sería imposible saber si éstas se debieron a la vacuna o al efecto placebo. Para eliminar esta confusión, todos los participantes habían creído que fueron tratados de la misma forma. Esto asegura que cualquier efecto placebo ocurriría igualmente en ambos grupos, de modo que los investigadores podrían atribuir cualquier diferencia restante a la vacuna.

Efectos del experimentador

Incluso si los sujetos de estudio no saben si recibieron el tratamiento real o el placebo, los *experimentadores* aún pueden tener un efecto. Por ejemplo, en la prueba de un medicamento antidepresivo, los experimentadores quizás entrevistarán a los pacientes para determinar si se sienten mejor. Pero si los experimentadores saben quién recibió el tratamiento y quién el placebo, sin

A propósito...

El efecto placebo puede ser extraordinariamente poderoso. En algunos estudios hasta 75% de los participantes que recibieron el placebo mejoraron en realidad. Para algunos pacientes, el efecto es tan poderoso que suplican continuar con su tratamiento aun después que se les dice que se les dio un placebo en lugar del tratamiento real. Sin embargo, otros investigadores no están de acuerdo en el poder del efecto placebo, e incluso algunos cuestionan la realidad del efecto.

A propósito...

Un efecto relacionado, conocido como el *efecto Hawthorne*, ocurre cuando los sujetos tratados, por alguna razón, responden de manera diferente sólo porque son parte de un experimento, sin importar la forma particular en la que son tratados. El efecto Hawthorne debe su nombre al hecho de que se observó por primera vez en un estudio de obreros en la fábrica Western Electric's Hawthorne.

darse cuenta podrían sonreír más a las personas en el grupo de tratamiento. Sus sonrisas podrían mejorar los humores de los participantes, haciendo parecer que el tratamiento funcionó, cuando de hecho la mejora fue causada por el experimentador. Este tipo de confusión, en la que el experimentador de alguna manera influye en los resultados, se denomina **efecto del experimentador** (o efecto Rosenthal). La única manera de evitar estos efectos es asegurarse que el experimentador no sepa cuáles sujetos pertenecen a cada grupo.

Definición

Un **efecto del experimentador** ocurre cuando un investigador o experimentador influye de alguna manera en los sujetos a través de factores, tales como expresiones faciales, tono de voz o actitud.

EJEMPLO 8 ¿Abuso infantil?

En un caso famoso dos parejas de Bakersfield, California, fueron condenadas por abusar de docenas de niños en edad preescolar en su guardería. La evidencia del abuso resultó principalmente de las entrevistas con los niños. Sin embargo, la condena fue rechazada después de que un hombre había cumplido 14 años en prisión, cuando un juez reexaminó las entrevistas y concluyó que los niños habían dado las respuestas que los entrevistadores querían oír. Si consideramos a los entrevistadores como los experimentadores, es un ejemplo de un efecto del experimentador, ya que los entrevistadores influyeron las respuestas de los niños por medio del tono y estilo de su interrogatorio.

A propósito...

Muchos casos de supuesto abuso extendido en guarderías y preescolares están siendo revisados para ver si efectos del experimentador (quienes entrevistaron a los niños) pueden llevar a condenas erróneas. Afirmaciones similares de efectos del experimentador se han hecho en casos que involucran memoria reprimida, en la que el apoyo psicológico supuestamente ayuda a la gente a recuperar recuerdos perdidos de eventos traumáticos.



Prueba ciega

En terminología estadística, la práctica de mantener a la gente sin saber quién está en el grupo de tratamiento y quién está en el grupo de control se denomina **prueba ciega**. Un experimento **ciego simple** es aquél en el que los participantes no saben a cuál grupo pertenecen, pero los experimentadores sí lo saben. Si ni los participantes ni los experimentadores saben quién pertenece a cada grupo, el estudio se dice que es **doble ciego**. Por supuesto, *alguien* conserva el control de los dos grupos para evaluar los resultados al final. Por lo común, en un experimento doble ciego los investigadores llevan a cabo el estudio contratando experimentadores para hacer cualquier contacto que se necesite con los participantes. De este modo, los investigadores evitan cualquier contacto con los participantes, asegurándose que ellos no puedan influirlos de alguna manera. El estudio de la vacuna contra la polio de Salk fue doble ciego, ya que ni los participantes (los niños) ni los experimentadores (los doctores y las enfermeras que aplicaron las inyecciones y realizaron el diagnóstico de polio) conocían a quiénes se les aplicó la vacuna y a quiénes se les administró el placebo.

Pruebas ciegas en experimentos

Un experimento es **ciego simple** si los participantes no saben si están en el grupo de tratamiento o en el grupo de control, pero los experimentadores sí lo saben.

Un experimento es **doble ciego** si ni los participantes ni los experimentadores saben quién pertenece al grupo de tratamiento y quién pertenece al grupo de control.

EJEMPLO 9 ¿Qué está mal con este experimento?

Para cada uno de los experimentos que se describen a continuación, identifique los problemas y explique cómo pueden evitarse.

- Se supone que un medicamento nuevo para el trastorno de déficit de atención (TDA) provoca que los niños sean más amables. Niños que sufren TDA fueron seleccionados aleatoriamente y divididos en grupos de tratamiento y de control. El experimento es ciego simple. Los experimentadores evalúan cuán amables son los niños durante una entrevista de uno en uno.

- b. Investigadores de la educación se preguntan si escuchar música clásica cuando se estudia mejora el aprendizaje. Ellos dieron a dos grupos de estudiantes una lección de dos horas y luego les permitieron tiempo para estudiar para un examen breve. Un grupo, compuesto de 50 estudiantes quienes les dijeron a los investigadores que les gustaba la música clásica, escuchó música clásica mientras estudiaba. El otro grupo, compuesto de 50 estudiantes quienes dijeron que no les gustaba la música clásica, estudió en silencio. Los resultados muestran que los estudiantes que escucharon música clásica hicieron mejor el examen.
- c. Investigadores buscan saber si los efectos de una rara enfermedad degenerativa puede detenerse mediante ejercicio. Identifican a 6 personas que padecen la enfermedad y de forma aleatoria asignan 3 al grupo de tratamiento, que hace ejercicio diario, y asignan a 3 al grupo de control que no hace ejercicio. Después de seis meses comparan la cantidad de degeneración en cada grupo.
- d. Un quiropráctico realiza ajustes en 25 pacientes con dolor de espalda. Después, 18 de los pacientes dicen que se sintieron mejor. Él concluye que los ajustes son un tratamiento efectivo.

Solución

- a. Los experimentadores valoran la cortesía en entrevistas, pero ya que saben cuáles niños recibieron el medicamento real, de manera inadvertida podrían hablar de forma diferente a estos niños durante las entrevistas. O podrían interpretar el comportamiento de los niños de manera diferente, ya que ellos saben cuáles sujetos recibieron el medicamento real. Éstos son efectos del experimentador que pueden confundir los resultados del estudio. El experimento debe ser doble ciego.
- b. El problema con este estudio es que los estudiantes no fueron asignados a los dos grupos al azar. Al colocar a los estudiantes que les gusta la música clásica en un grupo y a los que no les gusta en otro, los investigadores crearon una situación en la cual los dos grupos no comparten las mismas características generales.
- c. Los resultados del estudio serían difíciles de interpretar ya que los tamaños de las muestras no son suficientemente grandes. Por supuesto, con una enfermedad poco común, puede ser difícil encontrar a gente que participe en un experimento.
- d. Los 25 pacientes que recibieron ajustes representan un grupo de tratamiento, pero el estudio carece de un grupo de control. Los pacientes podrían haberse sentido bien a consecuencia del efecto placebo, en lugar de que sea por el efecto real de los ajustes. El quiropráctico podría haber mejorado su estudio si hubiera contratado a un actor que hiciese ajustes falsos (uno que se sintiera de manera parecida, pero que en realidad no se realizase de acuerdo con las guías quiroprácticas) en un grupo de control. Luego podría haber comparado los resultados en los dos grupos para ver si un efecto placebo estaba presente.

EJEMPLO 10 Identificación del tipo de estudio

Para cada una de las preguntas siguientes, ¿qué tipo de investigación estadística es más probable que conduzca a una respuesta?

- a. ¿Cuál es el ingreso medio de un corredor de la bolsa?
- b. ¿Los cinturones de seguridad salvan vidas?
- c. ¿El levantamiento de pesas mejora los tiempos de los corredores de la carrera de 10 km (10 K)?
- d. ¿El contacto de la piel con cierto pegamento causa sarpullido?
- e. ¿Puede un remedio de hierbas reducir la gravedad de los resfriados?
- f. ¿Complementos alimenticios de resveratrol (un extracto de uvas) prolongan la vida?

Solución

- a. Un *estudio de observación* puede decirnos el ingreso medio de un corredor de la bolsa. Sólo necesitamos encuestar a los corredores y la encuesta misma no cambiará los ingresos.

Con un tratamiento adecuado, un resfriado puede curarse en una semana. No le haga caso y puede durar siete días.

—Refrán médico popular

A propósito...

Experimentos con ratones muestran que éstos viven más y tienen mayor resistencia cuando toman dosis grandes de resveratrol, un extracto de la cáscara de uvas rojas. Sin embargo, aún no se sabe si dosis altas de resveratrol tendrían los mismos efectos en humanos, ni se sabe si el resveratrol es seguro para los humanos. Además, las dosis dadas a los ratones son tan grandes que podría ser poco práctico dar una dosis equivalente a los humanos, por nuestro tamaño tan grande, comparado con el de los roedores.



- b. Podría ser poco ético hacer un experimento en que se le pidiese a unas personas usar el cinturón de seguridad y a otros decirles que *no* lo usen. Por tanto, un estudio para determinar si los cinturones de seguridad salvan vidas debe ser *de observación*. Puesto que algunas personas eligen utilizar los cinturones de seguridad mientras que otros eligen no usarlos, podríamos llevar a cabo un *estudio en retrospectiva*. Al comparar la tasa de mortalidad en accidentes entre los que usaron y los que no utilizaron cinturones de seguridad, podemos aprender si los cinturones de seguridad salvan vidas. (Sí las salvan).
- c. Necesitamos un *experimento* para determinar si levantar pesas puede mejorar los tiempos de los corredores de 10 K. Al azar seleccionamos un grupo de corredores para crear un grupo de tratamiento al que sometemos a un programa de levantamiento de pesas y un grupo de control a quienes les pedimos que permanezcan alejados de las pesas. Debemos de tratar de asegurar que todos los demás aspectos de su entrenamiento sean similares. Luego veremos si los corredores en el grupo de levantamiento mejoran sus tiempo más que los del grupo de control. No podemos utilizar una prueba ciega en este experimento, ya que no hay manera de evitar que los participantes sepan que están levantando pesas.
- d. Un *experimento* puede ayudarnos a determinar si el contacto de la piel con el pegamento causa un sarpullido. En este caso es mejor utilizar un *experimento ciego sencillo* en el que aplicamos el pegamento real a los participantes en un grupo y aplicamos un placebo que parezca igual, pero que carezca del ingrediente activo, a los miembros del grupo de control. No hay necesidad de un experimento doble ciego ya que parece poco probable que los experimentadores puedan influir para que una persona tenga sarpullido. (Sin embargo, la pregunta de si el sujeto *tiene* un sarpullido está sujeta a interpretación, el conocimiento del experimentador de a quienes les dio el tratamiento real podría afectar esta interpretación).
- e. Debemos utilizar un *experimento doble ciego* para determinar si un nuevo remedio con base en hierbas puede reducir la gravedad de los resfriados. Algunos participantes reciben el remedio real, mientras que a otros se les da un placebo. Necesitamos las condiciones de doble ciego, ya que la severidad de un resfriado podría ser afectada por el humor u otros factores en que los investigadores, sin darse cuenta, podrían influir.
- f. El resveratrol ha sido identificado y se ha puesto disponible en forma de complementos sólo recientemente, por lo que necesitamos muchos años de datos para determinar si tiene el efecto de prolongar la vida. Por tanto, debemos utilizar un estudio en prospectiva diseñado para monitorear a participantes durante muchos años. Los participantes podrían conservar registros si han tomado, y cuánto, del complemento, y en algún momento los investigadores podrían analizar los datos para ver si el resveratrol tiene el efecto de prolongar la vida.

Meta-análisis

Todas las investigaciones estadísticas individuales son estudios de observación o experimentos. Sin embargo, en años recientes, los estadísticos han encontrado útil “extraer” grupos de estudios pasados para ver si pueden aprender algo que no fueron capaces de aprender de los estudios individuales. Por ejemplo, cientos de estudios habían considerado los efectos posibles de la vitamina C sobre los resfriados, por lo que los investigadores podrían revisar los datos de muchos de estos estudios como un grupo. Este tipo de estudio, en los que los investigadores revisan estudios anteriores como un grupo, se denomina **meta-análisis**.

Definición

En un **meta-análisis** los investigadores revisan muchos estudios anteriores. El meta-análisis considera a estos estudios como un grupo combinado, con el objetivo de encontrar tendencias que no fueron evidentes en los estudios individuales.

ESTUDIO DE CASO**Medicamentos para combatir la depresión:
un meta-análisis**

Los investigadores del gobierno estiman que 1 de cada 5 estadounidenses sufren de depresión en algún momento de su vida y que el costo anual para la economía en pérdida de productividad es al menos de 40 mil millones de dólares. Hasta la década de los noventa, los únicos medicamentos disponibles para combatir la depresión pertenecían a la clase conocida como antidepresivos tricíclicos. Pero en la década pasada, varios medicamentos se hicieron de uso común. El más famoso fue el Prozac, que había sido prescrito de manera tan amplia que algunas personas aseguraban que ahora en Estados Unidos se vivía en el “país del Prozac”. Pero, ¿el Prozac y los otros medicamentos sirven? Para responder a esta pregunta, los investigadores de la Agencia para las Políticas e Investigación del Cuidado de la Salud llevaron a cabo un meta-análisis.

Los investigadores empezaron por buscar en la literatura médica estudios acerca del tratamiento de la depresión. Encontraron más de 8 000 estudios reportados durante un periodo de 9 años. Al buscar estudios que cumplieren ciertos criterios especiales, tal como observación de los pacientes al menos durante seis semanas y la comparación de los tipos anterior y nuevo de medicamentos antidepresivos, redujeron la lista a casi 300 estudios. También consideraron alrededor de 600 estudios que trataban con efectos laterales de los medicamentos. Su meta-análisis consistió en analizar los resultados de todos estos estudios como un grupo, para buscar tendencias que podrían no haber sido evidentes en los estudios individuales.

Globalmente encontraron que los nuevos medicamentos son más efectivos que los placebos para personas gravemente deprimidas. Alrededor de 50% de tales personas respondieron a los medicamentos nuevos, comparadas con casi 32% que respondieron a un placebo. Sin embargo, los antidepresivos tricíclicos anteriores fueron igualmente efectivos y, con frecuencia, disponibles a precios menores. Los medicamentos nuevos y los anteriores tienen efectos secundarios diferentes, pero los grados de gravedad de estos efectos fueron aproximadamente iguales. Aunque estos resultados pueden considerarse como buenas y malas noticias para los fabricantes y usuarios de los medicamentos nuevos, el meta-análisis quizá fue más valioso por lo que podría *no* decir. Los investigadores encontraron que los datos fueron inadecuados para determinar si los medicamentos nuevos eran efectivos para depresión ligera o para niños, a pesar de que eran comúnmente recetados en ambos casos. También encontraron que los datos eran inadecuados para evaluar tratamientos con hierbas, tal como kava o mosto de San Juan. Así, el meta-análisis señaló la dirección en que era necesaria una nueva investigación.

UN MOMENTO DE REFLEXIÓN

Vea de nuevo los porcentajes de respuesta a los medicamentos nuevos (50%) y al placebo (32%) en el meta-análisis de medicamentos antidepresivos. ¿Las respuestas relativamente grandes al placebo sugieren que a los pacientes se les debe dar un placebo antes de suministrarles un medicamento real con efectos laterales potenciales? ¿Por qué sí o por qué no? Si usted fuera psicólogo, ¿qué le sugeriría a alguien que fue parte del 50%, que no respondió al medicamento real?

Sección 1.3 Ejercicios**Alfabetización estadística
y pensamiento crítico**

- 1. Placebo.** ¿Qué es un placebo y por qué es importante en un experimento para probar la efectividad de un medicamento?
- 2. Pruebas ciegas.** ¿Qué es una prueba ciega y por qué es importante en un experimento para probar la efectividad de un medicamento?
- 3. Confusión.** Al probar la efectividad de una vacuna nueva, suponga que los investigadores utilizaron hombres para el grupo de tratamiento y mujeres para el grupo placebo. ¿Cuál es la confusión y cómo afecta a este experimento?
- 4. Tratamiento negado.** En un estudio tristemente célebre, a 399 hombres afroamericanos con sífilis *no* se les dio un tratamiento que podría haberlos curado. A otros se les dieron tratamientos. La intención fue aprender acerca del efecto de la sífilis en hombres afroamericanos. ¿Este es un experimento o un estudio de observación? ¿Por qué este estudio es moral y terriblemente erróneo?
- 5. Color de ropa.** Un investigador planea indagar acerca de la creencia de que la gente se siente más a gusto en el sol de

verano, cuando utiliza ropa con colores claros, en lugar de ropa con colores oscuros. En este caso, ¿tiene sentido utilizar una prueba doble ciego?, ¿es sencillo implementar una prueba ciega en este caso?

6. **Tratamiento de césped.** Un investigador planea probar la efectividad de un nuevo fertilizante en el crecimiento de pasto. En este caso, ¿tiene sentido utilizar una prueba doble ciego?
7. **Medición del dolor.** Muchos estudios incluyen métodos para medir el dolor después de algún tratamiento. Describa cómo el efecto de un tratamiento podría afectar de manera adversa y describa cómo este efecto puede evitarse.
8. **Tratamiento de la calvicie.** En un estudio real de un tratamiento para la calvicie, se aseguró que el grupo placebo tenía mejores resultados en el crecimiento de nuevo cabello que el grupo tratado con el tratamiento. ¿Es posible este resultado? Estos resultados, ¿qué sugerirían acerca del tratamiento?

Conceptos y aplicaciones

Tipo de estudio. Para los ejercicios del 9 al 20, establezca el tipo de estudio e identifique las variables de interés. Sea tan específico como pueda.

9. **Terapia de toque.** Emily Rosa, una niña de nueve años, se convirtió en protagonista de un artículo en la *Journal of the American Medical Association* después de que ella probó el toque de terapeutas profesionales (vea la sección “Hablemos de educación” en el capítulo 10). Por medio de una división con un cartón ella mantenía una de sus manos encima de una de las manos del terapeuta, y se le pedía identificar la mano que Emily eligió.
10. **Control de calidad.** La Administración Federal de Medicamentos selecciona aleatoriamente una muestra de tabletas de aspirinas Bayer. La cantidad de aspirina en cada tableta se mide para ver su exactitud.
11. **Brazaletes magnéticos.** A los pasajeros de algunos cruces los proporcionan un brazalete magnético, y ellos acceden a utilizarlo en un intento por eliminar o disminuir los efectos del malestar por el movimiento. A otros se les proporcionan brazaletes similares que no tienen magnetismo.
12. **Selección de género.** En un estudio del método para la selección de género YSORT, desarrollado por el Instituto Genetics & IVF, 152 parejas tienen 127 niños y 25 niñas.
13. **Gemelos.** Un estudio de cientos de mellizos suecos determinó que el nivel de habilidad mental era más similar en gemelos idénticos (gemelos que provienen de un solo óvulo) que en gemelos fraternos (gemelos que provienen de dos óvulos distintos) (*Science*).
14. **Cardiopatía.** Un estudio europeo a 1 500 hombres y mujeres con niveles excepcionalmente altos del aminoácido homocisteína determinó que estos individuos tenían el doble de riesgo de una cardiopatía. Sin embargo, el riesgo era mucho menor para aquellos que tomaron complementos de vitamina B (*Journal of the American Medical Association*).
15. **Cáncer de próstata.** Un análisis de 11 estudios individuales intentó determinar si hay una relación concluyente entre la vasectomía y la incidencia de cáncer en la próstata (*Chance*).
16. **Encuesta de AOL.** Una encuesta de America OnLine (AOL) tuvo como resultado 38 410 respuestas a la pregunta “¿Cuánta credibilidad le da a pronósticos de clima de largo plazo?”. Entre aquellos que respondieron, 47% dijo “muy poca” o “ninguna”.
17. **Maíz OGM.** Investigadores de la Universidad de Nueva York encontraron que el maíz modificado genéticamente (OGM) conocido como maíz Bt libera un insecticida a través de sus raíces en la tierra (*Nature*).
18. **Investigación de célula madre.** En una encuesta de *Newsweek*, realizada por Princeton Survey Research Associates, 48% de los 1002 adultos que respondieron dijeron que estaban a favor de utilizar fondos federales para financiar investigación médica que utilice células madre de embriones.
19. **Tratamiento con imanes.** En un estudio acerca de los efectos de imanes en el dolor de espalda, algunos sujetos fueron tratados con imanes mientras que a otros se les dieron dispositivos no magnéticos pero con una apariencia similar. Los imanes no parecieron ser efectivos en el tratamiento del dolor de espalda (*Journal of the American Medical Association*).
20. **Cables de tendido eléctrico y cáncer.** Cientos de estudios científicos y estadísticos se han hecho para determinar si los cables aéreos de tendido eléctrico de alto voltaje aumentan la incidencia de cáncer entre los que viven cerca de ellos. Un estudio sumario con base en muchos estudios previos concluyó que no existe relación significativa entre los cables de tendido eléctrico y el cáncer (*Journal of the American Medical Association*).

¿Qué está mal en este experimento? Para cada uno de los estudios descritos en los ejercicios 21 a 28, identifique problemas que probablemente causen confusión y explique cómo pueden evitarse. Analice cualquier otro problema que podría afectar los resultados.

21. **Crecimiento de álamos.** Un experimento diseñado para evaluar la efectividad del riego y el fertilizante en el crecimiento de álamos: el fertilizante se utiliza con un grupo de álamos en una región húmeda y la irrigación en álamos en una región árida.
22. **Compras en internet.** Doscientos voluntarios se reclutan por medio de un anuncio en *PC Magazine*. Ellos son utilizados en un estudio para comparar los montos totales gastados por aquellos que usan internet y por los que no. A cada persona se le permite elegir estar en el grupo de usuarios de internet o el grupo que no utiliza internet. Al cabo de un mes se comparan los montos gastados por los dos grupos.

- 23. Nivel de octano.** En una comparación de gasolina con diferentes niveles de octano, 24 camionetas son llenadas con gasolina de 87 octanos, mientras que 28 automóviles deportivos son llenadas con gasolina de 91 octanos. Después que un vehículo ha sido conducido durante 250 millas, se mide la cantidad de gasolina consumida.
- 24. Ensayo con aspirina.** En la fase I de un ensayo clínico para probar la efectividad de la aspirina para la prevención de infartos al miocardio, la aspirina se le da a 3 personas y un placebo a otras 7.
- 25. Ensayo con hipertensión.** En un ensayo clínico acerca de la efectividad de un medicamento para tratar la hipertensión, a los sujetos se les pregunta si desean recibir el medicamento real o un placebo.
- 26. Pie de atleta.** En un ensayo clínico acerca de la efectividad de una loción utilizada para tratar *tinea pedis* (pie de atleta), los médicos que evalúan los resultados saben a qué sujetos se les aplicó el tratamiento y a quiénes se les aplicó un placebo.
- 27. Levantamiento de pesas.** En una prueba acerca de los efectos de levantamiento de pesas en la presión sanguínea, un grupo experimenta un tratamiento, que consiste en un programa de levantamiento de pesas, mientras que otro grupo levanta pelotas de tenis.
- 28. Mezclas de pinturas.** En pruebas de durabilidad de dos mezclas de pinturas diferentes, los investigadores que evalúan los resultados saben cuáles muestras son de cada una de las dos diferentes mezclas.

Análisis de experimentos. Los ejercicios 29 al 32 presentan preguntas que podrían ser guías en un experimento. Si usted tuviese que diseñar el experimento, ¿cómo seleccionaría a los grupos de tratamiento y de control? ¿El experimento debe ser ciego sencillo, doble ciego o ninguno de ellos? Explique su razonamiento.

- 29. Beethoven e inteligencia.** ¿Escuchar a Beethoven hace más inteligentes a los niños?
- 30. Lipitor y colesterol.** ¿El medicamento Lipitor tiene como resultado disminuir los niveles de colesterol?
- 31. Etanol y millaje.** ¿Un aditivo de etanol en la gasolina reduce el millaje?
- 32. Puertas en casa.** ¿Las puertas de servicio de aluminio son más en una casa que las puertas de servicio de madera?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 1 en www.aw.com/bbt.

- 33. Efectos del experimentador en casos de memoria reprimida.** Busque en la web artículos e información acerca de la controversia en cuanto a los casos de recuperación de memoria reprimida. Brevemente resuma uno o dos de los casos

más interesantes y, con base en lo que lea, exprese su opinión acerca de si las memorias recuperadas presuntamente lo hicieron influidas por los efectos del experimentador.

- 34. Ética en experimentos.** Los estándares éticos cambian de época a época. Un caso de sobra poco ético fue un estudio de sífilis llevado a cabo en Tuskegee, Alabama, de 1932 a 1972. (Vea el ejercicio 4). En dicho estudio se les dijo a hombres afroamericanos que recibirían un tratamiento para la sífilis, pero en realidad no lo recibieron. El objetivo oculto de los investigadores era estudiar los resultados a largo plazo de los efectos de la enfermedad. Utilice la web para aprender acerca de la historia del estudio de la sífilis en Tuskegee. Mantenga un debate en clase acerca de los temas éticos incluidos en este caso o escriba un breve ensayo que resuma el caso y sus lecciones éticas.
- 35. Debate: ¿deben utilizarse datos de experimentos no éticos?** Con frecuencia investigaciones realizadas en el pasado no cumplen con los estándares éticos actuales. En casos extremos, tal como la investigación realizada por doctores en la Alemania nazi, los investigadores con frecuencia mataban a los sujetos de sus experimentos. Aunque esta investigación no ética claramente violó los derechos humanos de los sujetos experimentales, en algunos casos llevó a revelaciones que podrían ayudar a la gente en nuestros días. ¿Es ético utilizar los resultados de investigación no ética?
- 36. Estudios detenidos prematuramente.** En ocasiones sucede que un estudio se detiene antes de su conclusión. Utilice internet para encontrar un ejemplo de tal estudio. ¿Por qué el estudio se detuvo? ¿Debió detenerse o hubiese sido mejor completar el estudio?



EN LAS NOTICIAS



- 37. Estudios de observación.** Busque en los periódicos de hace unas semanas y encuentre un ejemplo de una investigación estadística que sea de observación. Describa brevemente el estudio y resuma sus conclusiones.
- 38. Estudio experimental.** Busque en los periódicos de hace unas semanas y encuentre un ejemplo de una investigación estadística que incluya un experimento. Describa brevemente el estudio y resuma sus conclusiones.
- 39. Estudios en retrospectiva.** Busque en los periódicos de hace unas semanas y encuentre un ejemplo de una investigación estadística que sea de observación y retrospectiva. Describa brevemente el estudio y resuma sus conclusiones.
- 40. Meta-análisis.** Busque en los periódicos de hace unas semanas y encuentre un ejemplo de un meta-análisis. Describa brevemente el estudio y resuma sus conclusiones.



1.4 ¿Debe tener confianza en una investigación estadística?

Si he hecho algún descubrimiento valioso, ha sido más por la atención paciente, que por cualquier otro talento.

—Isaac Newton

La mayor parte del resto de este libro está dedicada a ayudarle a obtener una comprensión más profunda de los conceptos y definiciones que hemos estudiado hasta aquí. Pero ya sabe suficiente para lograr uno de los objetivos principales de este texto: ser capaz de responder la pregunta “¿Debe tener confianza en una investigación estadística?”.

La mayoría de los investigadores llevan a cabo sus investigaciones estadísticas con honestidad e integridad, y la mayoría de ellas se realizan con diligencia y cuidado. Sin embargo, la investigación estadística es de tal complejidad que puede surgir sesgo de muchas formas diferentes, haciendo muy importante que siempre examinemos con cuidado los informes de investigaciones estadísticas. No hay una manera definitiva de responder la pregunta: “¿Debe tener confianza en una investigación estadística?”. Sin embargo, en esta sección consideramos ocho directrices que pueden ser útiles. A lo largo del camino también introduciremos unas cuantas definiciones y conceptos que lo prepararán para los análisis que vienen más adelante.

Ocho directrices para la evaluación crítica de una investigación estadística

1. Identifique el objetivo del estudio, la población considerada y el tipo de estudio.
2. Considere la fuente, en particular si los investigadores pueden predisponerse.
3. Examine el método de muestreo para decidir si es probable producir una muestra representativa.
4. Busque problemas en la definición o medición de las variables de interés, cuáles pueden hacer difícil la interpretación de los resultados reportados.
5. Tenga cuidado de las variables de confusión que puedan invalidar las conclusiones de un estudio.
6. Considere el contexto y la redacción de las preguntas de la encuesta, busque cualquier aspecto que pueda tener una tendencia a producir respuestas imprecisas o deshonestas.
7. Verifique que los resultados estén bien representados en gráficas y conclusiones, ya que tanto los investigadores como los medios con frecuencia crean gráficas engañosas o sacan conclusiones que los resultados no apoyan.
8. Tome distancia y considere las conclusiones. ¿El estudio logra sus objetivos? ¿Las conclusiones tienen sentido? ¿Los resultados tienen algún significado práctico?

Los reportes de las noticias no siempre proporcionan suficiente información para que aplique todas las directrices anteriores, pero por lo regular puede encontrar información adicional en la web. Busque pistas tales como “reportado por la NASA” o “publicado en *New England Journal of Medicine*” para ayudarle a seguir las fuentes originales u otra información relevante.

Directriz 1: identifique el objetivo, la población y el tipo de estudio

El primer paso en la evaluación de una investigación estadística es entender el objetivo y el enfoque del estudio. Con base en lo que oiga o lea acerca de un estudio, trate de responder estas preguntas básicas:

- ¿Para qué fue diseñado el estudio?
- ¿Cuál fue la población bajo estudio? ¿La población estaba clara y apropiadamente definida?
- ¿El estudio fue de observación, un experimento o un meta-análisis? Si fue un estudio de observación, ¿fue retrospectivo? Si fue un experimento, ¿fue simple o doble ciego, y los grupos de tratamiento y de control fueron aleatorizados correctamente? Dado el objetivo, ¿el tipo de estudio fue apropiado?

EJEMPLO 1 ¿Tipo de estudio apropiado?

Imagine el reporte siguiente (hipotético) en un periódico: “Investigadores dieron a 100 participantes sus horóscopos individuales y les preguntaron si los horóscopos fueron precisos. De los participantes, 85% dijeron que sus horóscopos fueron precisos. Los investigadores concluyeron que los horóscopos son válidos la mayor parte del tiempo”. Analice este estudio de acuerdo con la directriz 1.

Solución El objetivo del estudio era determinar la validez de los horóscopos. Con base en los reportes de las noticias, parece que el estudio fue *de observación*: los investigadores sólo preguntaron a los participantes acerca de la precisión de los horóscopos. Sin embargo, ya que la precisión de un horóscopo es un tanto subjetiva, el estudio debió haber sido un experimento controlado en el que a algunas personas les diesen sus horóscopos reales y a otros un horóscopo falso. Luego los investigadores podrían buscar diferencias entre los dos grupos. Además, como los investigadores podrían influir con facilidad en los resultados, en la forma en que ellos cuestionan a los participantes, el experimento debe ser doble ciego. En resumen, el tipo de estudio no fue adecuado para el objetivo y sus resultados carecen de sentido.

UN MOMENTO DE REFLEXIÓN

Intente probar su horóscopo. Encuentre el horóscopo de ayer para cada uno de los 12 signos y ponga cada uno en una hoja separada, sin nada que identifique el signo. Revuelva las hojas de manera aleatoria y pida a algunas personas que adivinen cuál se suponía que es su horóscopo personal. ¿Cuántas personas hicieron la elección correcta? Analice sus resultados.

EJEMPLO 2 ¿La aspirina previene ataques cardíacos?

Un estudio reportado en *New England Journal of Medicine* (vol. 318, núm. 4) buscaba determinar si la aspirina es efectiva en la prevención de ataques cardíacos. Incluía 22 000 hombres que los médicos consideraron que estaban en riesgo de ataques cardíacos. Los hombres fueron divididos en un grupo de tratamiento que tomó aspirina y un grupo de control que no lo hizo. El resultado fue tan convincente en favor de los beneficios de la aspirina que, por razones éticas, el experimento se detuvo antes de que se completase y los sujetos fueron informados de los resultados. Muchos artículos salieron con la noticia titular que tomar aspirina puede ayudar a prevenir ataques cardíacos. Analice este titular de acuerdo con la directriz 1.

Solución El estudio fue un experimento, que es apropiado, y sus resultados parecen ser convincentes. Sin embargo, el hecho de que la muestra sólo consistió de hombres significa que los resultados deben considerarse para aplicación en la población de hombres. Puesto que los resultados de pruebas médicas en hombres no necesariamente se aplican a mujeres, las noticias tergiversaron los resultados cuando no cualificaron a la población.

Directriz 2: considere la fuente

Las investigaciones estadísticas cuentan con el supuesto de que son objetivas, pero la gente que las lleva a cabo y que las financia puede predisponerse. Por tanto, es importante considerar la fuente de un estudio y evaluar la posibilidad de sesgo que podría invalidar las conclusiones del estudio.

El sesgo puede ser obvio en casos donde una investigación estadística se lleva a cabo para marketing, promoción u otros propósitos comerciales. Por ejemplo, un anuncio de dentífrico que asegura que “4 de 5 dentistas prefieren su marca”, parece que tiene bases estadísticas, pero no nos dan detalles de cómo se llevó a cabo la encuesta. Puesto que, obviamente, los anunciantes quieren decir cosas buenas de sus marcas, es difícil tomar con seriedad la afirmación estadística sin más información acerca de cómo se obtuvieron los resultados.

Otros casos de sesgo pueden ser más sutiles. Por ejemplo, suponga que un estudio llevado a cabo de manera cuidadosa concluye que un medicamento nuevo ayuda a curar el cáncer. En apariencia, el estudio podría parecer creíble. Pero, ¿qué pasa si el estudio fue financiado por

A propósito...

Las encuestas revelan que casi la mitad de los estadounidenses creen en sus horóscopos. Sin embargo, en experimentos controlados, las predicciones de los horóscopos son verdaderas con menos frecuencia que lo que se esperaría por el azar.



A propósito...

Muchos estudios recientes muestran diferencias sustanciales en las formas en que hombres y mujeres responden a los mismos tratamientos médicos. Por ejemplo, la aspirina es más efectiva en sangre ligera en los hombres que en mujeres (se piensa que la sangre ligera ayuda a prevenir ataques cardíacos en algunas personas). La morfina controla el dolor mejor en mujeres, pero el ibuprofeno es más efectivo para hombres. Y las mujeres rechazan los trasplantes de corazón con mayor frecuencia que los hombres. Estas diferencias pueden derivarse de interacciones con las hormonas que difieren en hombres y mujeres o de diferencias en la velocidad en que los hombres y las mujeres metabolizan medicamentos diferentes.

una compañía farmacéutica que ganará miles de millones de dólares en ventas si el medicamento prueba ser efectivo? Los investigadores podrían haber realizado bien su trabajo con gran integridad a pesar de la fuente de financiamiento, pero podría ser valioso un poco de investigación adicional para estar seguros.

Las principales investigaciones estadísticas por lo regular son evaluadas por expertos imparciales. Por ejemplo, el proceso mediante el cual los científicos examinan investigaciones de otros se denomina **revisión de pares** (ya que los científicos que hacen la evaluación son *pares* de quienes la realizaron). Revistas científicas acreditadas requieren que todos los reportes de investigación sean revisados por pares antes que la investigación sea aceptada para su publicación. La revisión de pares no garantiza que un estudio sea válido, pero da credibilidad, ya que implica que otros expertos están de acuerdo en que el estudio fue llevado a cabo de manera adecuada.

Definición

Revisión de pares es un proceso en el que varios expertos en un campo evalúan un reporte de investigación antes de que el reporte sea publicado.

A propósito...

Después de décadas de argumentar lo contrario, en octubre de 1999 la compañía Philip Morris, el mayor vendedor de productos de tabaco, reconoció públicamente que fumar causa cáncer en los pulmones, cardiopatías, enfisema y otras enfermedades graves. A partir de esto, Philip Morris cambió su nombre a Altria.

EJEMPLO 3 ¿Es saludable fumar?

En 1963 las investigaciones habían mostrado tan claramente los peligros para la salud de fumar, que el director general de salud pública de Estados Unidos anunció públicamente que fumar es malo para la salud. Las investigaciones realizadas desde entonces dieron más apoyo a esta afirmación. Sin embargo, mientras que una gran mayoría de los estudios mostraron que fumar no es saludable, unos cuantos estudios no encontraron peligros en el fumar e incluso quizá *beneficios* para la salud. Por lo general, estos estudios fueron llevados a cabo por el Instituto de Investigación del Tabaco, financiado por las compañías tabacaleras. Analice estos estudios de acuerdo con la directriz 2.

Solución Incluso en un caso como éste puede ser difícil decidir a quién creer. Sin embargo, los estudios que muestran que fumar no es saludable aparecieron principalmente en investigaciones revisadas por pares. En contraste, los estudios realizados en el Instituto de Investigaciones de Tabaco tenían un claro sesgo potencial. El sesgo *potencial* no significa que la investigación fue sesgada, pero el hecho que contradiga a casi todas las otras investigaciones sobre el tema debe causar preocupación.

EJEMPLO 4 Conferencia de prensa científica

Suponga que las noticias nocturnas de televisión presentan a científicos en una rueda de prensa anunciando que han descubierto evidencia de que un químico nuevo desarrollado puede detener el proceso de envejecimiento. El trabajo aún no ha pasado por un proceso de revisión de pares. Analice el estudio de acuerdo con la directriz 2.

Solución Con frecuencia los científicos anuncian los resultados de sus investigaciones en una conferencia de prensa, por lo que el público puede enterarse de su trabajo tan pronto como sea posible. Sin embargo, podría requerirse mucha experiencia para evaluar su estudio sobre posible sesgo u otros errores, que es el objetivo del proceso de revisión por pares. Hasta que el trabajo sea revisado por pares y publicado en una revista de prestigio, cualesquiera hallazgos deben considerarse preliminares, en especial acerca de afirmaciones sorprendentes tal como la posibilidad de detener el proceso de envejecimiento.

Directriz 3: examine el método de muestreo

Una investigación estadística no puede ser válida a menos que la muestra sea representativa de la población bajo estudio. Métodos de muestreo malos casi garantizan una muestra sesgada que hace que los resultados del estudio sean inservibles.

Muestras sesgadas pueden surgir de muchas formas, pero dos problemas estrechamente relacionados son particularmente comunes. El primer problema, denominado **sesgo de selección** (o **efecto de selección**), ocurre siempre que los investigadores *seleccionan* su muestra de forma que los datos tiendan a no ser representativos de la población. Por ejemplo, un sondeo previo a elecciones que encuesta sólo a republicanos registrados tiene sesgo de selección, ya que es poco probable que reflejen las opiniones de los votantes de todos los partidos (e independientes).

El segundo problema, denominado **sesgo de participación**, puede surgir cuando la gente *elige* ser parte de un estudio, esto es, cuando los participantes son voluntarios. La forma más común de sesgo de participación ocurre en **encuestas de autoselección** (o **encuestas de respuesta voluntaria**), encuestas o sondeos en los que la gente decide participar. En tales casos, cuando la gente siente mucha atracción por un tema es más probable que participe, y sus opiniones podrían no representar las opiniones de la población en su totalidad, que tienen menos atracción sentimental hacia el tema.

Definición

El **sesgo de selección** (o **efecto de selección**) ocurre siempre que los investigadores *seleccionan* su muestra de una manera sesgada.

El **sesgo de participación** ocurre en cualquier momento en un estudio que es voluntario.

Una **encuesta de autoselección** (o **encuesta de respuesta voluntaria**) es una en la cual la gente decide si ella estará incluida en la encuesta.

ESTUDIO DE CASO

La encuesta de la revista *Literary Digest* de 1936

La popular revista *Literary Digest* de la década de los treinta, predijo de manera exitosa los resultados de varias elecciones por medio de encuestas grandes. En 1936, los editores de *Literary Digest* llevaron a cabo una encuesta grande previa a la elección presidencial. De manera aleatoria seleccionaron una muestra de 10 millones de personas de varias listas, incluyendo nombres del directorio telefónico y listas de clubes de campo. Enviaron por correo una tarjeta de “voto” a cada uno de estas 10 millones de personas. Alrededor de 2.4 millones de personas regresaron sus tarjetas de voto. Con base en los votos regresados, los editores de *Literary Digest* predijeron que Alf Landon ganaría la presidencia por un margen de 57% a 43% sobre Franklin Roosevelt. En lugar de eso, Roosevelt ganó con 62% del voto popular. ¿Cómo fue que esa encuesta grande estuvo tan errónea?

La muestra sufrió tanto de sesgo de selección como de sesgo de participantes. El sesgo de selección surgió porque *Literary Digest* eligió sus 10 millones de nombres de manera que favorecían la elección de personas de clase alta. Por ejemplo, seleccionaron nombres de guías telefónicas, lo que significó elegir sólo aquellos que podían permitirse teléfono en 1936. De forma similar, los miembros de un club de campo por lo común son muy pudientes. El sesgo de selección también se inclinaba a favor del republicano Landon porque la afluencia de votantes de 1930 se inclinaba por los candidatos republicanos.

El surgimiento del sesgo de participación de quienes regresaron las papeletas por correo fue voluntario, por lo que la gente que creyó firmemente en la elección fue más propensa a estar entre quienes regresaron sus votos. El sesgo también se inclinaba a favor de Landon ya que él era el aspirante, la gente que no quería a Roosevelt como presidente podía expresar su deseo para cambiar regresando las papeletas. Juntas, las dos formas de sesgo hicieron inservibles los resultados de la muestra, a pesar del gran número de personas encuestadas.

EJEMPLO 5 Encuesta de autoselección

El programa de televisión *Nightline* llevó a cabo una encuesta en la que a los televidentes se les preguntó si las oficinas de las Naciones Unidas deberían continuar en Estados Unidos. Los televidentes podrían responder la encuesta por un pago de 50 centavos al llamar a un número “900” con sus opiniones. La encuesta obtuvo 186 000 respuestas, de las cuales 67% estuvo a

A propósito...

Un encuestador joven llamado George Gallup llevó a cabo su propia encuesta antes de la elección de 1936. Él envió tarjetas a sólo 3 000 personas elegidas aleatoriamente y predijo no sólo el resultado de la elección, sino también el resultado del sondeo de *Literary Digest* dentro de 1%. Gallup continuó fundando una muy exitosa compañía de encuestas.

A propósito...

Más de un tercio de todos los estadounidenses habitualmente cierran la puerta o cuelgan el teléfono cuando son contactados para una encuesta, esto hace la autoselección un problema para los encuestadores legítimos. Una razón por la cual la gente cuelga puede ser la proliferación de ventas bajo la apariencia de una investigación de mercado (con frecuencia llamada “insinuación”) en la que el vendedor por teléfono pretende ser parte de una encuesta para tratar de que usted compre algo.

favor de sacar a las Naciones Unidas de Estados Unidos. En la misma época, un sondeo que utilizó muestreo aleatorio simple a 500 personas determinó que 72% quería que las Naciones Unidas *permanecieran* en Estados Unidos. ¿Cuál encuesta es más probable que sea representativa de la opinión general de los estadounidenses?

Solución La muestra de *Nightline* estuvo muy sesgada. Tuvo sesgo porque su muestra sólo se extrajo de los televidentes del programa, en lugar de todos los estadounidenses. La encuesta misma fue una encuesta de autoselección en la que los televidentes no sólo elegían responder, sino también tenían que *pagar* 50 centavos por participar. Este costo hizo aún más probable que los encuestados fuesen aquellos que sentían una necesidad de cambio. Por tanto, a pesar de su número tan grande de encuestados, la encuesta de *Nightline* es poco probable que diese resultados significativos. En contraste, el muestreo aleatorio simple de 500 personas es muy probable que sea representativa, de modo que los resultados de esta encuesta tienen mejor oportunidad de representar la opinión verdadera de todos los estadounidenses.

Apropósito...

La misión *Kepler* de la NASA programó un lanzamiento a finales de 2008, es un telescopio orbitando que debe ser capaz de detectar planetas tan pequeños como la Tierra alrededor de otras estrellas. *Kepler* buscará ligeras disminuciones de la luz de una estrella cada vez que un planeta que orbita pase enfrente suyo (un “tránsito”), lo cual significa que sólo será capaz de detectar planetas para la pequeña fracción de estrellas que tengan a sus sistemas planetarios alineados con nuestra línea visual.



EJEMPLO 6 Planetas alrededor de otras estrellas

Hasta mediados de la década de los noventa, los astrónomos no habían encontrado evidencia concluyente de planetas fuera de nuestro sistema solar. Pero la tecnología mejorada hizo posible empezar a buscar tales planetas y más de 200 se han descubierto hasta principios de 2007. La tecnología existente hizo más sencillo encontrar planetas grandes que planetas pequeños y más sencillo encontrar planetas que orbitan cerca de sus estrellas que planetas que orbitan lejos de sus estrellas. De acuerdo con la teoría principal de la formación de un sistema solar, los planetas que orbitan cerca de sus estrellas deben ser muy raros. Pero ellos fueron muy comunes entre los primeros 200 descubrimientos. ¿Esto significa que algo está mal con la teoría principal de la formación de un sistema solar?

Solución Aunque la teoría sugiere que planetas grandes en órbitas cercanas deben ser raros, la tecnología hace de estos casos raros los más fáciles de encontrar. Por tanto, la determinación de muchos de estos planetas representa un *efecto de selección* que sesga la muestra (de descubrimiento de planetas) hacia un tipo raro. De hecho, esto parece que es casi seguramente el caso, ya que actualmente muchos planetas “normales” se están descubriendo y los astrónomos han determinado que con pequeñas modificaciones la teoría existente puede explicar los tipos raros encontrados en los primeros descubrimientos.

Directriz 4: buscar problemas en la definición o medición de las variables de interés

Los resultados de una investigación estadística pueden ser difíciles de interpretar si las variables bajo estudio son difíciles de definir o medir. Por ejemplo, imagine tratar de llevar a cabo un estudio de cómo el ejercicio afecta el ritmo cardíaco en reposo. Las variables de interés serían *cantidad de ejercicio* y *ritmo cardíaco en reposo*. Ambas variables son difíciles de definir y medir. En el caso de *cantidad de ejercicio*, no es claro qué incluye la definición, ¿caminar está incluido? Incluso si especificamos la definición, ¿cómo medimos *cantidad de ejercicio* dado que algunas formas son más intensas que otras?

UN MOMENTO DE REFLEXIÓN

¿Cómo mediría su ritmo cardíaco en reposo? Describa algunas dificultades para definir y medir el ritmo cardíaco en reposo.

EJEMPLO 7 ¿El dinero puede comprar el amor?

Una encuesta de Roper reportada en *USA Today* incluyó una encuesta al 1% más ricos de los estadounidenses. La encuesta determinó que esta gente pagaría un promedio de \$487 000 por *amor verdadero*, \$407 000 por *gran inteligencia*, \$285 000 por *talento* y \$259 000 por *juventud eterna*. Analice estos resultados de acuerdo con la directriz 4.

Solución Las variables de dicho estudio son complejas de definir. Por ejemplo, ¿cómo definiría *amor verdadero*? y ¿qué significa amor verdadero por un día, toda la vida o algo semejante? De manera análoga, ¿la habilidad para equilibrar una cuchara en su nariz constituye *talento*? Puesto que las variables están tan vagamente definidas, es probable que personas diferentes las interpreten de manera distinta, haciendo que los resultados sean difíciles de interpretar.

EJEMPLO 8 Suministro ilegal de medicamento

Una estadística citada comúnmente es que las autoridades aplican la ley de manera exitosa para detener en sólo 10 a 20% la entrada ilegal de drogas a Estados Unidos. ¿Debe creer esta estadística?

Solución En esencia hay dos variables en un estudio de intercepción de droga ilegal: *cantidad de drogas ilegales interceptadas* y *cantidad de drogas ilegales NO interceptadas*. Debe ser relativamente sencillo medir la cantidad de drogas ilegales que los oficiales que aplican la ley interceptan. Sin embargo, puesto que las drogas son ilegales, es poco probable que alguien esté reportando la cantidad de droga que *no* es interceptada. Entonces, ¿cómo alguien puede saber que las drogas interceptadas son 10% a 20% del total? En un análisis del *New York Times*, un oficial de la policía fue cuestionado cuando dijo que sus colegas se referían a este tipo de estadísticas como “PFA” por sus siglas en inglés “sacadas del aire”.

Directriz 5: tener cuidado con las variables de confusión

Con frecuencia las variables que *no están planeadas* como parte del estudio pueden hacer difícil la interpretación adecuada de los resultados. Como se analizó en la sección 1.3, no siempre es sencillo descubrir estas *variables de confusión*. En ocasiones se descubren sólo años después de que un estudio se terminó y otras veces nunca se descubren, en cuyo caso la conclusión de un estudio podría aceptarse aunque no sea correcto. Por fortuna, las variables de confusión son más obvias y pueden descubrirse con sólo pensar detenidamente acerca de los factores que pueden influir en los resultados de un estudio.

EJEMPLO 9 Radón y cáncer de pulmón

El radón es un gas radiactivo producido por procesos naturales (decaimiento de uranio) en la Tierra. El gas puede filtrarse en los edificios por los cimientos y puede acumularse en concentraciones relativamente altas, si las puertas y ventanas están cerradas. Imagine un estudio (hipotético) que busca determinar si el gas radón causa cáncer en los pulmones comparando la tasa de cáncer de pulmón en Colorado, donde el gas es muy común, con la tasa de cáncer de pulmón en Hong Kong, donde el gas radón es menos común. Suponga que el estudio determina que las tasas de cáncer de pulmón son iguales. ¿Sería razonable concluir que el radón *no* es una causa importante de cáncer de pulmón?

Solución Las variables de interés son *cantidad de radón* (una variable de explicación en este caso) y *tasa de cáncer de pulmón* (variable de respuesta). Sin embargo, el gas radón no es la única causa posible de este cáncer. Por ejemplo, fumar puede causar cáncer pulmonar, por lo que en este estudio, *tasa de fumadores* podría ser una variable de confusión, en especial porque la tasa de fumadores en Hong Kong es mucho más alta que la de Colorado. Como resultado, no podemos sacar ninguna conclusión acerca de radón y cáncer pulmonar sin tomar en cuenta la tasa de fumadores (y quizás otras variables también). De hecho, estudios cuidadosos han mostrado que el gas radón *puede* causar cáncer pulmonar y la Agencia de Protección del Ambiente de Estados Unidos (EPA) recomendó seguir pasos para prevenir el radón en construcciones interiores.

A propósito...

Muchas ferreterías venden equipos sencillos que pueden utilizarse para probar si el gas radón está acumulándose en su casa. Si es el caso, el problema puede eliminarse instalando un sistema apropiado de “reducción de radón”, que por lo regular consiste en un ventilador que dispersa el radón debajo de la casa antes de que entre a ella.



Directriz 6: considerar el contexto y redacción de las encuestas

Aunque una encuesta se lleve a cabo de manera adecuada con un muestreo apropiado y con términos y preguntas claramente definidos, debe tener cuidado en la composición y palabras que podrían producir respuestas imprecisas o deshonestas. Estas últimas son particularmente probables cuando la encuesta tiene que ver con temas delicados, tales como hábitos personales o ingresos. Por ejemplo, la pregunta “¿Usted miente en su declaración de impuestos?” es poco probable que obtenga respuestas honestas de aquellos que mienten, a menos que estén completamente seguros de la confidencialidad (y quizá ni aun así).

En otros casos, incluso respuestas honestas podrían no ser realmente precisas, si la redacción del cuestionario invita al sesgo. Algunas veces el orden de las palabras en una pregunta puede afectar el resultado. Un sondeo realizado en Alemania hizo las dos preguntas siguientes:

- ¿Diría que el tráfico contribuye más o menos a la contaminación del aire que las industrias?
- ¿Diría que las industrias contribuyen más o menos a la contaminación del aire que el tráfico?

La única diferencia es el orden de las palabras *tráfico* e *industrias*, pero esta diferencia cambia drásticamente los resultados: con la primera pregunta 45% respondieron que el tráfico y 32% respondieron que las industrias. Con la segunda, sólo 24% respondió que el tráfico, mientras 57% respondió que las industrias.

A propósito...

En una encuesta es más probable que la gente seleccione los elementos que aparecen primero, a consecuencia de lo que los psicólogos llaman el *error de disponibilidad*, la tendencia a realizar juicios con base en lo que está a *disponibilidad* en la mente. Organizaciones profesionales de encuestas tienen mucho cuidado para evitar este problema; por ejemplo, ellos pueden plantear una pregunta con dos opciones en un orden para la mitad de las personas en la muestra y en orden inverso para la otra mitad.

EJEMPLO 10 ¿Quiere una disminución de impuestos?

En un tiempo, cuando el gobierno de Estados Unidos tuvo excedentes en el presupuesto anual, los republicanos en el congreso propusieron una disminución de impuestos y el Comité Nacional Republicano encargó una encuesta para determinar si los estadounidenses apoyaban la propuesta. Preguntaron, “¿Está a favor de una reducción de impuestos?”, 67% de los encuestados respondieron *sí*. ¿Debemos concluir que los estadounidenses apoyaban la propuesta?

Solución Una pregunta como “¿Está a favor de una reducción de impuestos?” está sesgada ya que no da otras opciones. De hecho, un sondeo independiente realizado al mismo tiempo daba a los encuestados una lista de opiniones para el uso de los ingresos excedentes. Esta encuesta encontró que 31% quería el dinero dedicado a Seguridad Social, 26% quería que se utilizara para reducir la deuda nacional y sólo 18% estaba a favor de usarlo para una reducción de impuestos. (El restante 25% de encuestados eligió entre una variedad de otras opciones).

EJEMPLO 11 Encuesta sensible

Dos encuestas preguntaron a católicos en el área de Boston si los anticonceptivos debían estar disponibles a mujeres solteras. La primera incluyó entrevistas personales y 44% de los encuestados respondieron que *sí*. La segunda encuesta fue realizada por correo y por teléfono y 75% de los encuestados respondieron *sí*.

Solución Los anticonceptivos son un tema delicado, en particular entre católicos (ya que la Iglesia católica oficialmente se opone a los anticonceptivos). La primera encuesta, con entrevistas personales podría promover respuestas deshonestas. La segunda encuesta hacía que las respuestas pareciesen más privadas y, por tanto, era más probable que refleje las opiniones verdaderas de los encuestados.

Directriz 7: verificar que los resultados estén bien representados en gráficas y conclusiones

Aun cuando una investigación estadística esté bien realizada, puede representarse de manera incorrecta en gráficas o conclusiones. Los investigadores en ocasiones malinterpretan los resultados de sus propios estudios o pasan a conclusiones que no están sustentadas por los resultados, en particular cuando tienen un sesgo personal. Reporteros de noticias podrían malinterpretar una encuesta o pasar a conclusiones no garantizadas que hacen que una historia parezca más

espectacular. Las gráficas engañosas son especialmente comunes (en el capítulo 3 dedicamos gran parte a este tema). Siempre debe buscar inconsistencias entre la interpretación de un estudio (en figuras y en palabras) y cualesquiera datos reales dados con él.

EJEMPLO 12 ¿El consejo escolar necesita una lección de estadística?

El consejo escolar en Boulder, Colorado, causó revuelo cuando anunció que 28% de los niños en escuelas de Boulder leían “por debajo del nivel de su grado” y por eso concluyeron que los métodos de enseñanza de lectura necesitaban ser cambiados. El anuncio estuvo basado en las pruebas de lectura en que 28% de los niños en edad escolar obtuvieron calificaciones por debajo del promedio nacional para su grado. ¿Estos datos apoyan la conclusión del consejo?

Solución El hecho de que 28% de los niños en edad escolar obtuvieran calificaciones por debajo del promedio nacional para su grado implica que 72% sacaron calificaciones en y por arriba del promedio nacional. Así, la ominosa declaración acerca de la lectura de los estudiantes “debajo del nivel de su grado” tiene sentido sólo si “nivel de su grado” significa la calificación promedio nacional para un grado particular. Esta interpretación de “nivel de su grado” es curiosa, ya que implicaría que la mitad de los estudiantes en el país siempre están debajo del nivel de su grado, no importa cuán altas sean las calificaciones. Aun podría ser el caso que los métodos de enseñanza necesitasen ser mejorados, pero estos datos no justificaban esa conclusión.

Directriz 8: reflexione y considere las conclusiones

Por último, incluso si un estudio parece razonable de acuerdo con todas las directrices anteriores, debe reflexionar y considerar las conclusiones. Hágase preguntas tal como éstas:

- ¿El estudio llegó a sus objetivos?
- ¿Las conclusiones tienen sentido?
- ¿Puede descartar explicaciones alternas para los resultados?
- Si las conclusiones tienen sentido, ¿tienen significado práctico?

EJEMPLO 13 Afirmaciones extraordinarias

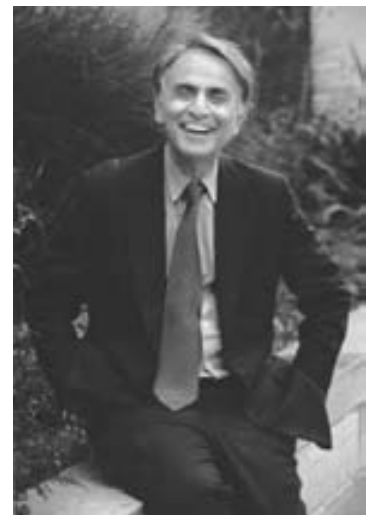
Suponga que un estudio (hipotético) concluye que usar una cadena de oro aumenta su probabilidad de sobrevivir en un accidente automovilístico en 10%. Esta afirmación tiene como base un análisis estadístico acerca de datos de tasas de supervivencia y lo que la gente llevaba puesto. Un análisis cuidadoso de la investigación muestra que fue realizada de manera adecuada y cuidadosa. ¿Usted debe llevar puesta una cadena de oro siempre que conduzca un automóvil?

Solución A pesar del cuidado del estudio, la afirmación de que una cadena de oro puede salvar su vida en un choque es difícil de creer. Después de todo, ¿cómo podría una delgada cadena de oro ayudar en un choque a alta velocidad? Ciertamente, es *posible* que algún efecto desconocido de las cadenas de oro hagan correcta la conclusión, pero parece mucho más probable que los resultados fueran debido a la casualidad o debido a una variable de confusión no identificada (por ejemplo, quizás aquellos que usan cadenas de oro son más pudientes y conducen automóviles más nuevos con mayores características de seguridad, disminuyendo la tasa de casos fatales).

EJEMPLO 14 Importancia práctica

Un experimento conducido en el que la pérdida de peso de una persona que consume un nuevo “complemento rápido dietético” es comparada con los pesos perdidos en un grupo de control que trata de perder peso de otras formas. Al cabo de ocho semanas, los resultados muestran que el grupo de tratamiento pierde un promedio de media libra más que el grupo de control. Suponiendo que no hay efectos secundarios, ¿este estudio sugiere que el complemento rápido dietético es un buen tratamiento para la gente que quiere bajar de peso?

Solución Comparada con el peso promedio de una persona, una pérdida de media libra difícilmente importa mucho. Así, aunque los resultados de pérdida de peso podrían ser interesantes, no parecen tener mucha importancia práctica.



Afirmaciones extraordinarias requieren de evidencia extraordinaria.

—Carl Sagan, astrónomo y ganador del Premio Pulitzer

Sección 1.4 Ejercicios

Alfabetización estadística y pensamiento crítico

- Revisión de pares.** ¿Qué es una revisión de pares? ¿Cómo es útil?
- Sesgo de selección y sesgo de participación.** Describa y compare el sesgo de selección y el sesgo de participación en muestreo.
- Encuestas de autoselección.** ¿Por qué las encuestas de autoselección siempre son propensas al sesgo de participación?
- Variables de confusión.** ¿Qué son las variables de confusión y qué problemas pueden causar?

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Validez de la encuesta.** La encuesta de televisión obtuvo más de un millón de respuestas por teléfono, por lo que es claramente más válida que la encuesta realizada por encuestadores profesionales, que incluyó entrevistas con sólo unos cientos de personas.
- Ubicación de encuestas.** La encuesta de creencias religiosas sufrió de sesgo de selección, ya que los cuestionarios sólo fueron repartidos en iglesias católicas.
- Vitamina C y resfriados.** Mi experimento probó, más allá de cualquier duda, que la vitamina C puede reducir la gravedad de los resfriados, puesto que controlé el experimento cuidadosamente para cualquier variable de confusión posible.
- Trote rápido.** Cualquiera que trote por ejercicio debe probar el nuevo régimen de entrenamiento, ya que estudios cuidadosos sugieren que puede aumentar su velocidad en 1%.

Conceptos y aplicaciones

Aplicación de directrices. En los ejercicios 9 al 16, determine cuál de las ocho directrices para evaluar estudios estadísticos parece ser más relevante. Explique su razonamiento.

- Control de calidad.** El departamento de relaciones públicas de Teletronics llevó a cabo un estudio de la tasa de defectos de los circuitos fabricados por la compañía.
- Selección de género.** Después de un método de selección de género diseñado para aumentar la verosimilitud de tener niña, se encontró que 23 bebés fueron niñas y 21 fueron niños, por lo que el portavoz de la compañía estableció que el método es efectivo la mayoría de las veces.
- Ética.** En un estudio a 250 abogados, a cada uno se le preguntó si él o ella tienen ética correcta.

- Agricultura.** Investigadores concluyeron que un sistema de riego usado para plantíos de tomates en California es más efectivo que un sistema competidor usado en Arizona.
- Encuesta a prisioneros.** Los prisioneros que asisten a clases a la universidad de uno de los autores fueron encuestados acerca de reincidencia y 10% de ellos respondió.
- Sondeo de elección.** Bajo el encabezado “Turner predijo ganar en forma aplastante”, se informó que Turner recibió 55% de los votos en una encuesta previa a la elección, comparada con 45% de su oponente.
- Encuesta de energía nuclear.** A adultos seleccionados de manera aleatoria se les hizo la pregunta: “¿Está de acuerdo o en desacuerdo con el aumento de la producción de energía nuclear que, potencialmente, podría causar la muerte de miles de personas inocentes?”.
- Bienes falsificados.** Un consorcio de fabricantes planea un estudio diseñado para comparar el valor de bienes falsificados en Estados Unidos en el año 2000 con el valor de los producidos en el año actual.

Sesgo. En los ejercicios 17 a 20, identifique y explique al menos una fuente de sesgo en cada estudio que se describe. Luego sugiera cómo podría evitarse el sesgo.

- Chocolate saludable.** El *New York Times* publicó un artículo que incluía estos enunciados: “Finalmente, el chocolate se mueve hacia su legítimo lugar en la pirámide de alimentos, en la parte más alta junto con el vino tinto, frutas, verduras y té verde. Varios estudios reportados en el *Journal of Nutrition* mostraron que después de comer chocolate, los sujetos de prueba habían aumentado los niveles de antioxidante en su sangre. El chocolate contiene flavonoides antioxidantes que se han asociado con la disminución de riesgo de cardiopatías y apoplejía. Mars Inc., la compañía de dulces y la Asociación de Fabricantes de Chocolate financiaron gran parte de la investigación.”
- Libro famoso.** Cuando la autora Shere Hite escribió *Women and Love: A Cultural Revolution in Progress*, basó sus conclusiones acerca de la población general de todas las mujeres en 4500 respuestas que recibió después de enviar 100 000 cuestionarios a varios grupos de mujeres.
- Napster.** La revista *Newsweek* efectuó una encuesta acerca del sitio web de Napster para descargar música. Los lectores podían registrar sus respuestas en el sitio web de *Newsweek*.
- Método de encuesta.** Usted planea llevar a cabo una encuesta para determinar el porcentaje de personas en su estado que pueden nombrar al vicegobernador (que planea competir por el senado en Estados Unidos). Usted obtiene direcciones de una lista de habitantes en el estado y envía por correo una encuesta a 850 personas seleccionadas aleatoriamente de la lista.

21. Todo está en la redacción. Princeton Survey Research Associates hizo un estudio para la revista *Newsweek* que ilustra el efecto de la redacción en una encuesta. Se hicieron dos preguntas:

- ¿Usted, personalmente, cree que el aborto está mal?
- Sin importar su punto de vista personal sobre el aborto, ¿usted está a favor o en contra de que una mujer en este país tenga la opción de tener un aborto con la ayuda de su doctora?

Para la primera pregunta, 57% de los encuestados respondieron *sí*, mientras que 36% respondieron *no*. En respuesta a la segunda pregunta, 69% de los encuestados estuvieron a favor de la opción, mientras que 24% se opuso. Analice por qué las dos preguntas producen resultados contradictorios. ¿Cómo podrían utilizarse los resultados de las preguntas de manera selectiva por varios grupos?

22. ¿Impuesto o gasto? Una encuesta de Gallup preguntó lo siguiente:

- ¿Usted está a favor de una reducción de impuestos o “aumento en el gasto en otros programas del gobierno”? *Resultado:* 75% para la disminución de impuestos.
- ¿Usted está a favor de una disminución de impuestos o “destinar dinero para financiar nuevas cuentas de ahorro para el retiro, así como aumentar el gasto en educación, defensa, seguro médico para ancianos y otros programas”? *Resultado:* 60% para el gasto.

Analice por qué las dos preguntas produjeron resultados contradictorios. ¿Cómo podrían utilizarse los resultados de las preguntas de manera selectiva por varios grupos?

¿Qué quiere conocer? Los ejercicios 23 a 26 plantean dos preguntas relacionadas que podrían formar la base de una investigación estadística. Brevemente analice en qué difieren las dos preguntas y cómo estas diferencias afectarían el objetivo de un estudio y el diseño del estudio.

23. Citas por internet

Primera pregunta: ¿Qué porcentaje de citas por internet terminan en matrimonio?

Segunda pregunta: ¿Qué porcentaje de matrimonios inician con citas por internet?

24. Profesores de tiempo completo

Primera pregunta: ¿Qué porcentaje de clases de primer año en un campus son enseñadas por los profesores de tiempo completo?

Segunda pregunta: ¿Qué porcentaje de profesores de tiempo completo dan clases de primer año?

25. Borracheras

Primera pregunta: ¿Con qué frecuencia los estudiantes universitarios se emborrachan?

Segunda pregunta: ¿Con qué frecuencia hay borracheras realizadas por estudiantes de universidad?

26. Cursos de estadística

Primera pregunta: ¿Cuál es la proporción de graduados de universidad que tomaron un curso de estadística?

Segunda pregunta: ¿Qué proporción de todos los cursos de estadística son tomados por estudiantes universitarios?

¿Encabezados precisos? Los ejercicios 27 y 28 dan un encabezado y una descripción breve de las estadísticas del artículo que acompañaron al encabezado. En cada caso analice si el encabezado representa de manera precisa al artículo.

27. Encabezado: “Las drogas se muestran en 98% de las películas”.

Resumen del artículo: un “estudio oficial” asegura que el uso de drogas, bebida o el fumar se describían en 98% de las películas más rentadas (Associated Press).

28. Encabezado: “El sexo es más importante que el trabajo”.

Resumen del artículo: una encuesta determinó que 82% de 500 personas entrevistadas por teléfono clasificaron una vida sexual satisfactoria como importante o muy importante, mientras que 79% clasificó trabajo satisfactorio como importante o muy importante (Associated Press).

Declaraciones políticas. Comúnmente los políticos creen que ellos deben hacer sus declaraciones políticas muy cortas ya que el periodo de atención de los oyentes es muy breve. Un efecto similar ocurre en reportes estadísticos en noticias. Las principales investigaciones estadísticas con frecuencia se reducen a una o dos oraciones. Los resúmenes de informes estadísticos en los ejercicios 29 a 32 se tomaron de varias fuentes de noticias. Describa cuál información crucial está perdida en el enunciado dado y qué más necesitaría conocer antes de seguir los consejos del informe.

29. Confianza en militares. Los informes de *USA Today* en una encuesta de Harris aseguran que el porcentaje de adultos con “gran confianza” en líderes militares se encuentra en 54% (arriba de 37% en 1997).

30. Mejores restaurantes. Informes de CNN en una encuesta de Zagat de los Mejores Restaurantes de Estados Unidos determinaron que “sólo nueve restaurantes lograron un estupendo 29 de una calificación de 30 posible y ninguno de esos restaurantes están en la Gran Manzana (Central Park en Nueva York)”.

31. Pronóstico de clima. Un encabezado de *USA Today* informó que “Más compañías intentan apostar a los pronósticos del clima”. El artículo dio ejemplos de compañías que creían que los pronósticos a largo plazo son creíbles y cuatro compañías fueron citadas.

32. Nacimientos en China. Un encabezado en *USA Today* informó que “China perdió el equilibrio cuando los niños superaron al número de niñas” y la gráfica adjunta mostró que por cada 100 niñas nacidas en China, nacieron 116.9 niños.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 1 en www.aw.com/bbt.

- 33. Análisis de una investigación estadística.** Encuentre un reporte detallado de alguna investigación estadística reciente de interés para usted. Escriba un informe breve aplicando cada uno de las ocho directrices dadas en esta sección. (Si algunas de las directrices no pueden aplicarse al estudio particular que esté analizando, explique por qué la directriz no es aplicable).
- 34. Índice de Harper.** Vaya al sitio web para el índice de Harper y examine algunas de las estadísticas recién citadas. Elija tres estadísticas que encuentre particularmente interesantes y analice si las cree, con base en las directrices dadas en esta sección.
- 35. Estudios de gemelos.** Investigadores que hacen estudios estadísticos en biología, psicología y sociología están agradecidos por la existencia de gemelos. Los gemelos pueden usarse para estudiar si ciertos rasgos son heredados de los padres (naturaleza) o adquiridos del entorno de uno durante la educación (aprendidos). Los gemelos idénticos están formados del mismo óvulo en la madre y tienen el mismo material genético. Los gemelos fraternos están formados de dos óvulos separados y comparten aproximadamente la mitad del mismo material genético. Encuentre un informe publicado de un estudio de gemelos. Analice cómo los gemelos idénticos y fraternos son utilizados para formar el grupo del caso y el de control. Aplique las directrices 1 a 8 para el estudio y comente si encuentra convincentes las conclusiones.
- 36. Revista *American Demographics*.** Consulte un tema de *American Demographics*, una revista no técnica que se especializa en resumir estudios estadísticos que incluye a

estadounidenses y sus estilos de vida. Seleccione un artículo específico y utilice las ideas de esta sección para resumir y evaluar el estudio. Asegúrese de citar la información que considere faltante y que debe proporcionarse para que usted haga un análisis completo.

EN LAS NOTICIAS

- 37. Aplicación de directrices.** Encuentre un artículo reciente en el periódico o en la televisión acerca de una investigación estadística sobre un tema que considere interesante. Escriba un informe breve que aplique cada una de las ocho directrices dadas en esta sección. (Algunas de las directrices podrían no aplicarse al estudio particular que está analizando, explique por qué la directriz no es aplicable).
- 38. Resultados creíbles.** Encuentre un informe de una noticia reciente acerca de una investigación estadística cuyos resultados usted considera significativos e importantes. En una página o menos, resuma el estudio y explique por qué lo considera creíble.
- 39. Resultados poco creíbles.** Encuentre un informe de una noticia reciente acerca de una investigación estadística cuyos resultados usted *no* considera significativos ni importantes. En una página o menos, resuma el estudio y explique por qué no cree sus afirmaciones.
- 40. Encuesta de autoselección.** Encuentre un ejemplo de una encuesta reciente en que la muestra fue de autoselección. Describa la conformación de la muestra y cómo cree que la autoselección afecta los resultados de la encuesta.

Ejercicios de repaso del capítulo

1. **Armas en casa.** Una encuesta de Gallup incluyó a adultos seleccionados en Estados Unidos. Se les hizo esta pregunta: “¿Usted tiene un arma en casa?”. Entre las 1 012 respuestas recibidas, 38% fueron *sí*. El margen de error reportado fue de 3 puntos porcentuales.
 - a. Interprete el margen de error identificando el rango de valores que es probable que contenga la proporción de todos los hogares con armas.
 - b. Identifique la población.
 - c. ¿El estudio es un experimento o un estudio de observación?
 - d. ¿El valor reportado de 38% es un parámetro poblacional o una estadística muestral? ¿Por qué?
 - e. Si sabe que los sujetos de la encuesta respondieron a un artículo de una revista que pidió a sus lectores sus respuestas por teléfono, ¿consideraría que los resultados de la encuesta son válidos? ¿Por qué sí o por qué no?
 - f. Describa un procedimiento para seleccionar los sujetos de la encuesta usando una muestra aleatoria simple.
 - g. Describa un procedimiento para seleccionar sujetos similares usando muestreo estratificado.
 - h. Describa un procedimiento para seleccionar sujetos similares usando muestreo por conglomerados.
 - i. Describa un procedimiento para seleccionar sujetos similares usando muestreo sistemático.
 - j. Describa un procedimiento para seleccionar sujetos similares usando muestreo de conveniencia.
2. **Muestra aleatoria simple.** Un elemento importante de este capítulo es el concepto de una muestra aleatoria simple.
 - a. ¿Qué es una muestra aleatoria simple?
 - b. Cuando la Oficina de Estadísticas Laborales realizó una encuesta, inició dividiendo la población adulta de Estados Unidos en 2 007 grupos denominados *unidades primarias de muestreo*. Suponga que estas unidades primarias de muestreo contienen el mismo número de adultos. Si selecciona aleatoriamente un adulto de cada unidad primaria de muestreo, ¿el resultado es una muestra aleatoria simple? ¿Por qué sí o por qué no?
 - c. Con respecto a las unidades de muestreo simple descritas en el inciso b, describa un plan de muestreo que resulte en una muestra aleatoria simple.
3. **Prueba de Zocor.** En ensayos clínicos del medicamento Zocor, usado para tratar altos niveles de colesterol, se observó a 1 583 usuarios de Zocor por reacciones adversas. Se encontró que 3.5% de ellos experimentaban dolor de cabeza.
 - a. Con base en la información dada, ¿puede concluir que en algunos casos el Zocor causa dolor de cabeza? ¿Por qué sí o por qué no?
 - b. Mientras que 1 583 sujetos fueron tratados con Zocor, a otros 157 sujetos se les dio un placebo y 5.1% del grupo placebo experimentaron dolor de cabeza. ¿Qué sugiere esta información adicional acerca de dolores de cabeza como una reacción adversa al uso de Zocor?
 - c. En este ensayo clínico, ¿qué es una prueba ciega? y ¿por qué es importante en la prueba de los efectos de Zocor?
 - d. Este ensayo clínico ¿es un estudio observacional o un experimento? Explique.
 - e. ¿Qué es un efecto del experimentador y cómo podría minimizarse este efecto?
4. **Redacción de una pregunta en una encuesta.** En *The Superpollsters*, David W. Moore describe un experimento en el que a sujetos diferentes se les preguntó si estaban de acuerdo con los enunciados siguientes:
 - Muy poco dinero se gasta en el bienestar.
 - Muy poco dinero se gasta en asistencia a los pobres.

Aunque son los pobres quienes reciben bienestar, sólo 19% estaban de acuerdo cuando la palabra *bienestar* era usada, pero 63% estaba de acuerdo con *asistencia a los pobres*.

 - a. ¿Cuál de las dos preguntas debe utilizarse en la encuesta? ¿Por qué?
 - b. Si usted estuviese trabajando en una campaña para un candidato conservador para el Congreso y necesitase enfatizar la oposición al uso de fondos federales para la asistencia a los pobres, ¿cuál de las dos preguntas usaría? ¿Por qué?
 - c. ¿Es ético redactar deliberadamente la pregunta de una encuesta de modo que influya en las respuestas? ¿Por qué sí o por qué no?

Cuestionario del capítulo

Elija la mejor respuesta para cada una de las preguntas siguientes. Explique su razonamiento con una o más oraciones completas.

1. Usted realiza una encuesta en la cual selecciona aleatoriamente a 1 000 votantes registrados en Texas y pregunta si aprueban el trabajo que su gobernador está haciendo. La población para este estudio es
 - a. todos los votantes registrados en el estado de Texas.
 - b. las 1 000 personas que entrevistó usted.
 - c. el gobernador de Texas.
2. Para la encuesta descrita en la pregunta 1, sería más probable que sus resultados sufrieran sesgo si eligió a los participantes de
 - a. todos los votantes registrados en Texas.
 - b. todas las personas con licencia de manejo en Texas.
 - c. personas que donaron dinero para la campaña del gobernador.
3. Cuando decimos que una muestra es *representativa* de la población, queremos decir que
 - a. los resultados encontrados para la muestra son similares a los que encontraríamos para toda la población.
 - b. la muestra es muy grande.
 - c. la muestra fue seleccionada de la mejor manera posible.
4. Considere un experimento diseñado para ver si incentivos en efectivo pueden mejorar la asistencia a la escuela. La investigadora selecciona dos grupos de 100 estudiantes de secundaria y ofrece a un grupo \$10 por cada semana de asistencia perfecta. Ella le dice al otro grupo que son parte de un experimento pero no les da incentivo alguno. Los estudiantes que no reciben incentivo representan
 - a. el grupo de tratamiento.
 - b. el grupo de control.
 - c. el grupo de observación.
5. El experimento descrito en la pregunta 4
 - a. es ciego sencillo.
 - b. es doble ciego.
 - c. no es ciego.
6. El propósito de un placebo es
 - a. evitar que los participantes conozcan si pertenecen al grupo de tratamiento o al grupo de control.
 - b. distinguir entre los casos y los controles en un estudio de control de caso.
 - c. determinar si enfermedades pueden curarse sin tratamiento alguno.
7. Si vemos un efecto placebo en un experimento para probar un tratamiento diseñado para curar verrugas
 - a. el experimento no fue apropiadamente doble ciego.
 - b. los grupos experimentales fueron demasiado pequeños.
 - c. las verrugas fueron curadas entre los miembros del grupo de control.
8. Un experimento es ciego sencillo si
 - a. carece de un grupo de tratamiento.
 - b. carece de un grupo de control.
 - c. los participantes no saben si pertenecen al grupo de tratamiento o al grupo de control.
9. La encuesta X predice que Powell recibirá 49% de los votos, mientras que la encuesta Y predice que ella recibirá 53% de los votos. Ambas encuestas tienen un margen de error de 3 puntos porcentuales. ¿Qué puede concluir?
 - a. Una de las encuestas debe haberse realizado de manera incorrecta.
 - b. Las dos encuestas son consistentes una con la otra.
 - c. Powell recibirá 51% de los votos.
10. Una encuesta revela que 12% de los estadounidenses creen que Elvis aún está vivo, con un margen de error de 4 puntos porcentuales. El intervalo para esta encuesta es
 - a. de 10% a 14%.
 - b. de 8% a 16%.
 - c. de 4% a 20%.
11. Un estudio realizado por Exxon Mobil muestra que no había daño permanente a consecuencia de un gran derrame en Alaska. La conclusión
 - a. definitivamente no es válida, ya que el estudio fue sesgado.
 - b. puede ser correcto, pero el sesgo potencial significa que usted debe revisar con cuidado cómo se llegó a la conclusión.
 - c. podría ser correcto si cae dentro del intervalo de confianza del estudio.

- 12.** El programa de televisión *American Idol* selecciona ganadores de votos dados por cualquiera que quiere votar. Esto significa que el ganador
- a.** es la persona que la mayoría de los estadounidenses quiere que gane.
 - b.** puede o no ser la persona que la mayoría de los estadounidenses quiere que gane, ya que el voto está sujeto a sesgo de participantes.
 - c.** puede o no ser la persona que la mayoría de los estadounidenses quiere que gane, ya que el voto debió haber sido doble ciego.
- 13.** Considere un experimento en el que usted mide los pesos de niños de 6 años de edad. En este estudio la variable de interés es
- a.** el tamaño de la muestra.
 - b.** los pesos de niños de 6 años de edad.
 - c.** las edades de los niños bajo estudio.
- 14.** Imagine una encuesta de personas elegidas al azar en la cual se encuentra que la gente que utiliza crema de protección solar tuvo *más* probabilidad de sufrir quemaduras debidas al sol el año pasado. ¿Qué explicación parece más probable para este resultado?
- a.** la crema de protección solar no sirve.
 - b.** la gente en el estudio utilizó crema de protección solar que estaba caduca.
 - c.** la gente que utiliza crema de protección solar es más probable que pase tiempo en el sol
- 15.** Si una investigación estadística se realiza de manera cuidadosa de todas las formas posibles, entonces
- a.** sus resultados deben ser correctos.
 - b.** podemos tener confianza en sus resultados, pero aún es posible que no sean correctos.
 - c.** decimos que el estudio es perfectamente sesgado.



¿Las guarderías generan niños agresivos?

Hace unas décadas se daba por hecho que las mamás se quedaban en casa para cuidar a sus hijos pequeños. Sin embargo, entre más mujeres trabajadoras hay, más niños pasan mayor tiempo en algún tipo de guardería. En 1975 menos de 40% de las mujeres con hijos menores de 6 años trabajaban medio o tiempo completo fuera de casa. En nuestros días más de 60% de las madres son empleadas. Este drástico incremento en el número de niños en guardería llevó a psicólogos y sociólogos al estudio de los efectos de las guarderías. En 1991, el Instituto Nacional de Salud Infantil y Desarrollo Humano (NICHD, por sus siglas en inglés) inició el estudio del Cuidado de Niños a Edad Temprana, hizo el seguimiento de más de 1 300 niños inscritos en guarderías en 10 diferentes localidades de investigación en Estados Unidos y

continuó su seguimiento hasta los 15 años. Observe que esto establecía un estudio *prospectivo* de observación (vea la sección 1.3).

El estudio ha sido realizado cuidadosamente por investigadores destacados, a pesar de ello los resultados han provocado cierta controversia. Como ejemplo considere este resultado reportado en 2001: los niños que pasan más tiempo en guarderías son significativamente más propensos a ser agresivos cuando crezcan. Las madres trabajadoras de todas partes se sienten culpables y preocupadas por las consecuencias de las guarderías en sus hijos y los medios de comunicación informaron que las guarderías generaban niños agresivos. Pero, ¿debemos creer esta afirmación? Hay muchas maneras de evaluarla, pero por facilidad usemos las ocho directrices proporcionadas en la sección 1.4 (página 34).

Directriz 1: El objetivo del estudio parece claro (aprender sobre los efectos de una guardería) así como el grupo de estudio (todos los niños en guarderías). No obstante, el tipo de estudio plantea un problema potencial. El estudio es de observación más que un experimento aleatorizado, ya que sería poco ético recomendar o no a los padres el uso de las guarderías. Por desgracia, aunque un estudio de observación ayude a establecer una relación (o *correlación*) entre variables, ésta no puede por sí misma establecer causa y efecto (vea el capítulo 7). En este caso, el estudio nos da evidencia de una relación entre guarderías y niños agresivos, pero no menciona si uno es causa del otro. El problema es que puesto que los padres *eligieron* si sus hijos estuvieran en guarderías o en casa, no podemos asegurar que los grupos muestral y de control sean verdaderamente comparables. Por ejemplo, quizá las madres que hacen “malabares” entre los niños y el trabajo tienden a estar más estresadas que las madres que no trabajan; que a su vez puede hacer a los niños de guardería más estresados que los que no están en ellas. O podría ser que las madres que tienen hijos inquietos con mayor probabilidad los metan a guarderías. De este modo, sin importar la validez de la relación entre guardería y agresividad, la naturaleza de observación del estudio hace imposible sostener la afirmación que la guardería *cause* la agresividad.

Directriz 2: El estudio se ha llevado cuidadosamente y no tenemos ninguna razón para sospechar algún sesgo por parte de los investigadores. Sin embargo, las guarderías es un tema de carga emocional y los investigadores podrían tener opiniones fuertes acerca de ellas, incrementando el sesgo potencial. En realidad, diferentes investigadores han interpretado los resultados del estudio de diferentes maneras, lo cual sugiere que el sesgo puede desempeñar un papel en las interpretaciones, pero no en el estudio mismo.

Directriz 3: Los investigadores utilizaron un tipo de muestreo estratificado para reclutar participantes para el estudio, ya que querían relacionar la demografía nacional con factores como nivel socio-económico, raza y estructura familiar. Este tipo de muestreo parece apropiado para el estudio. Sin embargo, las familias que fueron reclutadas tienen la opción de

participar en el estudio. Esta autoselección aumenta de manera automática el potencial sesgo de participación, que puede afectar los resultados del estudio.

Directriz 4: La afirmación sobre las guarderías y la agresividad incluye esencialmente dos variables: *el tiempo que un niño permanece en la guardería y el nivel de agresividad del niño*. Ambas son muy difíciles de definir. Por ejemplo, el estudio del Cuidado de Niños a Edad Temprana define *guardería* como cuidado proporcionado por alguna persona que no es la madre, esto significa que el cuidado de los padres o abuelos cuenta como guardería. Mucha gente impugna esta definición. Definir agresión también es complicado, por ejemplo, podría ser difícil distinguir entre un juego activo y verdadera agresividad.

Directriz 5: Hay muchas variables de confusión potenciales en el estudio del Cuidado de Niños a Edad Temprana. Como se vio al inicio, el estudio no proporciona información acerca de si las madres de los niños en guarderías son sujetas a más estrés que las mamás que permanecen en casa, lo cual podría confundir los resultados. Quizá la mayor preocupación es que no hay una forma aceptable para medir o controlar la calidad de las guarderías. En consecuencia, incluso algunos de los investigadores involucrados en el estudio han sugerido que la agresividad observada es un resultado de una guardería de pobre o mediocre calidad en comparación con las guarderías en general.

Directriz 6: Los investigadores utilizaron entrevistas para determinar el tipo y la cantidad de cuidado, así como el nivel de agresión de cada niño. Ya que la agresividad es difícil de definir y puede ser concebida de diferente manera por cada persona, no hay una forma clara para asegurar que todos los entrevistadores pregunten precisamente de la misma manera (no sólo en palabras, sino en expresión facial y entonación) y que todos los entrevistados interpreten las preguntas de la misma forma.

Directriz 7: Los informes producidos por los investigadores del Cuidado de Niños a Edad Temprana reconocen importantes fallas en el estudio y enfatizan que el estudio sólo muestra una *posible* relación entre guarderías y agresividad. Así que, ¿de dónde viene la afirmación de que las guarderías generan niños agresivos? Viene del modo en que los medios de comunicación presentan sus resultados. Ya que los actuales resultados no respaldan esta interpretación, los noticieros reflejan una clara falta de entendimiento sobre el tema. Por ejemplo, los periodistas que no han estudiado estadística no podrían distinguir entre los resultados de una relación y los resultados de una causa, y entonces su falta de entendimiento se refleja en sus noticias. De manera alterna, los medios de comunicación podrían deliberadamente dar un cariz sensacionalista a los resultados con tal de llamar más la atención. Y en algunos casos, personas que creen que las mujeres deben permanecer en casa con sus hijos, podrían tergiversar los resultados para afirmar que el estudio proporciona respaldo científico para sus opiniones.

Directriz 8: La pregunta sobre el significado práctico en este estudio es muy importante. A nivel individual: muchas mujeres trabajadoras no creen tener opción sobre las guarderías, ellas deben trabajar por razones financieras y profesionales. Por tanto, ellas podrían hacer uso de las guarderías aun si tuvieran pruebas de que el cuidado en casa fuera mejor. A nivel social, los resultados podrían ser prácticamente significativos si éstos nos condujeran a nuevas políticas que podrían mejorar el cuidado de los niños en general. Pero mientras que los resultados no prueben causa y efecto, no pueden dar una pauta clara para los que dictan la política.

En resumen, hemos encontrado que un título provocativo sobre las guarderías y agresividad se desmorona ante un examen riguroso. Aún así, la investigación podría ser muy valiosa, en especial porque los mismos investigadores reconocen importantes fallas en el estudio. Los resultados pueden llevar a futuros estudios a resultados más concluyentes, información que seguramente será apreciada por padres de familia trabajadores de todas partes. Pero hasta entonces necesitamos ser cuidadosos en cómo interpretar los resultados del estudio de Cuidado de Niños a Edad Temprana y más aún acerca de cómo reaccionamos ante títulos sensacionalistas que dicho estudio pudiera producir.

PREGUNTAS PARA DISCUSIÓN

1. Una tercera variable de confusión es *el tiempo que los niños pasan frente a la televisión o con videojuegos*, los cuales también han sido asociados con conducta agresiva. ¿Considera que esta variable pueda afectar los resultados del estudio del Cuidado de Niños a Edad

Temprana? y si es así, ¿cómo los afectaría? ¿Qué otras variables de confusión pueden afectar sus resultados? Explique claramente sus respuestas.

2. Suponga que estuviera diseñando una continuación del estudio con la intención de proporcionar resultados más concluyentes. Con base en lo que ha aprendido del estudio del Cuidado a Niños en Edad Temprana, ¿qué haría de manera diferente?
3. ¿Cuál es su opinión personal acerca de si es mejor para un padre permanecer en casa o trabajar? ¿Su opinión cambiaría si el estudio de Cuidado de Niños a Edad Temprana produjese algunos resultados definitivos acerca de las guarderías? ¿Por qué sí o por qué no?
4. En general, ¿considera que el estudio presentado fue un esfuerzo valioso en términos del tiempo y gasto que requirió? Defienda su opinión.

LECTURAS SUGERIDAS

Lea más acerca del Estudio del Cuidado de Niños a Edad Temprana (SECC, por sus siglas en inglés) en su sitio web en <http://secc.rti.org>.

Carey, Benedict, "Study Finds Rise in Behavior Problems After Significant Time in Day Care", *New York Times*, 26 de marzo de 2007.

Stolberg, Sheryl Gay, "Science, Studies, and Motherhood", *New York Times*, 22 de abril de 2001.

Sweeney, Jennifer Foote, "The Day-Care Scare, Again", *Salon.com*, 20 de abril de 2001.

HABLEMOS DE SALUD PÚBLICA

¿Es saludable su estilo de vida?

Considere las conclusiones siguientes de estudios estadísticos:

- Fumar aumenta el riesgo de cardiopatía.
- Comer margarina puede aumentar el riesgo de cardiopatía.
- Un vaso de vino diario puede proteger contra cardiopatía, pero aumenta el riesgo de cáncer de mama.

Quizás esté familiarizado con algunas de estas conclusiones y tal vez ha cambiado su estilo de vida por ellas. Pero, ¿de dónde provienen? Para su sorpresa éstas y cientos de otras conclusiones importantes sobre salud pública provienen de un enorme estudio prospectivo que ha proporcionado información para cientos de estudios estadísticos más pequeños. Conocido como el Estudio de Salud de Enfermeras de Harvard, es el estudio de salud pública más longevo jamás emprendido. Si éste aún no ha cambiado su forma de vida, casi seguramente lo hará en el futuro.

Dicho estudio inició en 1976 cuando el doctor Frank E. Speizer, un profesor en la Escuela de Medicina de Harvard, decidió estudiar los efectos a largo plazo de anticonceptivos orales. Él envió cuestionarios por correo a aproximadamente 370 000 enfermeras registradas y recibió más de 120 000 respuestas. Eligió encuestar a enfermeras pues creía que su entrenamiento médico haría que sus respuestas fuesen más creíbles que las del público.

Cuando el doctor Speizer y sus colegas examinaron cuidadosamente la información en los cuestionarios regresados, se dieron cuenta que el estudio podría ampliarse para incluir más cosas que sólo los efectos de los anticonceptivos. En nuestros días este equipo de investigación continúa siguiendo alrededor de 90% de las 120 000 encuestadas originalmente. Cientos de investigadores también utilizan la información para estudiar salud pública.

Los cuestionarios anuales aún son una parte vital del estudio, lo que permite a los investigadores reunir información acerca de lo que las enfermeras comen; qué medicinas y vitaminas toman; si se ejercitan y cuánto, toman y fuman; y qué enfermedades han contraído. Algunas de las enfermeras también proporcionan muestras de sangre, que es usada para medir cosas tales como nivel de colesterol, niveles hormonales, variaciones genéticas y residuos de pesticidas y contaminantes del medio ambiente. La confianza del doctor Speizer en las enfermeras ha comprobado estar justificada, ya que dependen de las encuestas y ellas casi siempre, a solicitud, proporcionan muestras de sangre extraídas y etiquetadas adecuadamente.

Después de más de 30 años de correspondencia, tanto los investigadores como las enfermeras sienten un sentido de cercanía. Muchas de las enfermeras quedan en espera de oír noticias de los investigadores y dicen que el estudio les ha ayudado a poner más atención en cómo viven sus vidas. Los investigadores sienten profunda pena cuando deben registrar la muerte de una de las enfermeras.

La pena por el fallecimiento desempeñará un papel creciente en el estudio, ya que muchas de ellas ahora están entrando a sus 70 y 80. Pero esta pena también dará una abundante información acerca de los factores que influyen en la longevidad y salud en edad avanzada. Los investigadores desean que la información marcará el camino hacia una comprensión definitiva acerca de lo que constituye una dieta saludable, cómo la contaminación y la exposición a químicos afectan la salud y cómo podríamos prevenir las enfermedades degenerativas como la osteoporosis y el Alzheimer. En muerte, las 120 000 mujeres del Estudio de Salud de Enfermeras de Harvard pueden dar el regalo de una mejor vida a las generaciones futuras.



PREGUNTAS PARA DISCUSIÓN

1. Considere algunos resultados que quizá vienen del Estudio de Salud de Enfermeras de Harvard durante los siguientes 10 a 20 años. ¿Qué tipos de resultados cree que serán más importantes? ¿Considera que las conclusiones alterarán la manera en que vive su vida?
2. Explique por qué este estudio es de observación. Los críticos algunas veces dicen que los resultados serían más válidos si se obtuvieran por experimentos en lugar de observaciones. Analice si sería posible reunir información similar llevando a cabo experimentos de una manera ética y práctica.
3. En principio, este estudio está sujeto a sesgo de participación ya que sólo 120 000 de los 370 000 cuestionarios originales fueron devueltos. ¿Los investigadores deben preocuparse por este sesgo? ¿Por qué sí o por qué no?
4. Otra falla potencial proviene del hecho de que los cuestionarios con frecuencia abordaban temas sensibles de salud personal y los investigadores no tenían forma de confirmar que las enfermeras respondían de manera honesta. ¿Considera que la deshonestidad podría llevar a los investigadores a conclusiones incorrectas? Defienda su opinión.
5. Todas las participantes en el Estudio de Salud de Enfermeras de Harvard son mujeres. ¿Considera que también los resultados son de uso para los hombres? ¿Por qué sí o por qué no?
6. Haga una búsqueda en la web por artículos que analicen los resultados de este estudio. Elija un resultado reciente que le interese y analice lo que significa y cómo puede afectar, en el futuro, la salud pública o su propia salud.

LECTURAS SUGERIDAS

Lea más acerca del Estudio de Salud de Enfermeras de Harvard (con sede en el Laboratorio Channing de Harvard) en su sitio web en <http://www.nursehealthstudy.org>.

Brophy, Beth, "Doing It for Science", *U.S. News & World Report*, vol. 126, 21 marzo de 1999, p. 67.

Conwell, Vikki, "Forever Friends: Ties Can Last a Lifetime and Sustain Body and Soul", *Atlanta Journal-Constitution*, 2 de enero de 2007.

Parker-Pope, Tara, "Drinking Has Hidden Health Risks for Women", *Wall Street Journal*, 26 de diciembre de 2006.

Yoon, Carol Kaesuk, "In Nurses' Lives, a Treasure Trove of Health Data", *New York Times*, 15 de septiembre de 1998.



Prácticamente nadie sabe de lo que está hablando cuando aparecen números en los periódicos. Y eso es porque siempre consultamos a otras personas quienes no saben de lo que están hablando, como políticos y analistas del mercado.

—Molly Ivins (1944-2007)
columnista de periódico

Medición en estadística

TODOS SABEMOS CÓMO MEDIR CANTIDADES TALES

como estatura, peso y temperatura. Sin embargo, en la investigación estadística existen muchas otras clases de medidas, y debemos asegurarnos que estén definidas y reportadas de manera cuidadosa. En este capítulo analizamos algunos conceptos importantes asociados con las medidas y la estadística. Como verá, estos conceptos son muy útiles para comprender los informes estadísticos que encuentra en su vida diaria.

OBJETIVOS DE APRENDIZAJE

2.1 Tipos de datos y niveles de medida

Ser capaz de identificar datos cualitativos o cuantitativos, para identificar los datos cuantitativos como discretos o continuos, y para asignar los datos de un nivel de medida (relación nominal, ordinal o intervalo).

2.2 Manejo de errores

Entender la diferencia entre los errores aleatorios y sistemáticos, ser capaz de describir los errores de su tamaño absoluto y relativo, y saber la diferencia entre exactitud y precisión en las mediciones.

2.3 Uso de porcentajes en estadística

Entender cómo los porcentajes se usan para informar los resultados estadísticos y reconocer las formas en que a veces son mal utilizados.

2.4 Números índice

Entender el concepto de un número de índice; en particular, entender cómo el Índice de Precios al Consumidor (IPC) se utiliza para medir la inflación.

2.1 Tipos de datos y niveles de medida

Uno de los retos en estadística es decidir cómo resumir y mostrar datos de la mejor manera. Diferentes tipos de datos requieren distintas formas de resúmenes. En esta sección analizaremos cómo se clasifican los datos, una idea que nos ayudará cuando consideremos resúmenes de datos y visualización de datos en los últimos capítulos.

Tipos de datos

Los datos son de dos tipos básicos: cualitativos y cuantitativos. Los **datos cualitativos** tienen valores que pueden colocarse en *categorías no numéricas*. (Por esta razón, en ocasiones a estos datos se les denomina datos *categoricos*). Por ejemplo, los datos acerca del color de los ojos son cualitativos, ya que están clasificados por colores, tales como azul, café y avellana. Otros ejemplos de datos cualitativos incluyen sabores de helado, nombres de empleadores, género de animales, y “número de estrellas” de películas o restaurantes. Observe que la clasificación por estrellas es cualitativa aunque incluya un *número* de estrellas (tal como tres estrellas o cuatro estrellas), ya que los números no son necesarios y no podrían utilizarse para cálculos; de igual forma podríamos clasificar las películas con cuatro categorías no numéricas, tal como mala, regular, buena y excelente.

Los **datos cuantitativos** tienen valores numéricos que representan cuentas o medidas. Los tiempos de los corredores en una carrera, los ingresos de graduados de una universidad y el número de estudiantes en clases diferentes, son ejemplos de datos cuantitativos.

Tipos de datos

Datos cualitativos (o **categoricos**) consisten en valores que pueden colocarse en categorías no numéricas.

Datos cuantitativos consisten en valores que representan cuentas o medidas.

EJEMPLO 1 Tipos de datos

Clasifique cada uno de los conjuntos de datos siguientes como cualitativos o cuantitativos.

- Nombres de marcas de zapatos en una encuesta a consumidores.
- Calificaciones en un examen de opción múltiple.
- Calificaciones con letras en una tarea de ensayo.
- Números en los uniformes que identifican a los jugadores en un equipo de básquetbol.

Solución

- Los nombres de marcas son categorías y, por tanto, representan datos cualitativos.
- Las calificaciones en un examen de opción múltiple son cuantitativas, ya que representan un conteo del número de respuestas correctas.
- Calificaciones con letras en una tarea de ensayo son cualitativas ya que representan diferentes categorías de desempeño (desde no aprobado hasta excelente).
- Los números en los uniformes de los jugadores son cualitativos ya que no representan un conteo o medida; sólo se utilizan para identificación. Podemos decir que estos números son cualitativos en lugar de cuantitativos, ya que no podríamos utilizarlos para hacer cálculos. Por ejemplo, no tendría sentido sumar o restar los números de los uniformes de jugadores distintos.

Datos discretos frente a datos continuos

Además, a los datos cuantitativos los clasificamos como continuos o discretos. Los datos son **continuos** si pueden tomar *cualquier* valor en un intervalo dado. Por ejemplo, el peso de una persona puede ser cualquiera entre 0 y unos cuantos cientos de libras, así, los datos que consisten en pesos son continuos. Los datos son **discretos** si sólo toman valores particulares y no otros valores entre ellos. Por ejemplo, el número de estudiantes en su clase es discreto, ya que debe ser un número entero no negativo, y la medida de los zapatos es discreta ya que sólo toma valores en los enteros o en mitad de enteros, tal como 7, $7\frac{1}{2}$, 8 y $8\frac{1}{2}$.

Datos discretos frente a datos continuos

Los datos **continuos** toman *cualquier* valor en un intervalo dado.

Los datos **discretos** sólo toman valores particulares, distintos y ningún valor entre ellos.

EJEMPLO 2 ¿Discreto o continuo?

Para cada conjunto de datos, indique si los datos son discretos o continuos.

- Medidas del tiempo que toma caminar una milla.
- Los números de años calendario (tal como 2007, 2008, 2009).
- Los números de vacas lecheras en diferentes granjas.
- La cantidad de leche producida por vacas lecheras en una granja.

Solución

- El tiempo toma cualquier valor, por lo cual las medidas de tiempo son continuas.
- Los números de años calendario son discretos ya que no pueden ser valores fraccionarios. Por ejemplo, en la víspera del año nuevo de 2009, el año cambió de 2009 a 2010; nunca diremos el año es $2009\frac{1}{2}$.
- Cada granja tiene un entero no negativo de vacas que podemos contar, por lo que estos datos son discretos. (Por ejemplo, no puede contar fracciones de vaca).
- La cantidad de leche que una vaca produce toma cualquier valor en algún rango, por lo cual la información de producción de leche es continua.

Niveles de medida

Otra forma de clasificar los datos es por su *nivel de medida*. El nivel más simple de medida se aplica a variables tal como color de ojos, sabores de un helado o género de animales. Estas variables son descritas por nombres, etiquetas o categorías. Decimos que los datos están en un **nivel de medida nominal**. (La palabra *nominal* indica *nombres* para las categorías). El nivel nominal de medida no incluye ninguna clasificación u ordenamiento de los datos. Por ejemplo, no podríamos decir que ojos azules están antes que los ojos color café o que la vainilla clasifica más alto que el chocolate.

Cuando describimos datos con un esquema de clasificación u orden, tal como número de estrellas de películas o restaurantes, estamos utilizando un **nivel de medida ordinal**. (La palabra *ordinal* se refiere a *orden*). Por lo común no podemos utilizar tales datos con algún sentido para cálculos. Por ejemplo, no tiene sentido sumar número de estrellas, ver tres películas de una estrella no es equivalente a ver una película de tres estrellas.

UN MOMENTO DE REFLEXIÓN

Considere una encuesta que pregunta: “¿Cuál es su sabor favorito de helado?”. Hemos dicho que los sabores de helado representan datos en el nivel nominal de medida. Pero suponga que, por conveniencia, los investigadores ingresan los datos de la encuesta en una computadora y asignan números a los diferentes sabores. Por ejemplo, asignan 1 = vainilla, 2 = chocolate, 3 = galletas y crema, etcétera. ¿Esto cambia los datos de sabor de helado de nominal a ordinal? ¿Por qué sí o por qué no?

El de medida nivel ordinal proporciona un sistema de clasificación, pero no permite determinar las diferencias precisas entre las medidas. Por ejemplo, no hay manera de determinar la diferencia exacta entre una película de tres estrellas y una de dos estrellas. En contraste, una temperatura de 81°F es más caliente que 80°F en la misma cantidad que 28°F es más caliente que 27°F . Los datos de temperatura están a un nivel más alto de medida, ya que los *intervalos* (diferencias) entre unidades en una escala de temperatura siempre significan la misma cantidad definida. Sin embargo, aunque los intervalos (que incluyen sustracción) entre temperaturas tienen significado, *razones* (que incluyen división) no. Por ejemplo, *no es cierto* que 20°F sea el doble de caliente que 10°F o que -40°F sea el doble de frío que -20°F . Por lo que las razones carecen de significado en la escala Fahrenheit, es que su *punto cero es arbitrario* y no representa un estado de “no calor”. Si los intervalos tienen significado pero las razones no, como en el caso de las temperaturas Fahrenheit, decimos que los datos están en el **nivel de medida de intervalo**.

Cuando tanto intervalos como razones tienen significado, decimos que los datos están en el **nivel de medida de razón**. Por ejemplo, datos que consisten en distancias están en el nivel de razón de medida, ya que una distancia de 10 km realmente está el doble de lejos que una distancia de 5 km. En general, el nivel de razón de medida se aplica a cualquier escala con un *cero verdadero*, que es un valor que significa *nada* de algo que está en medición. En el caso de distancias, una distancia de cero significa que “no hay distancia”. Otros ejemplos de datos al nivel de razón de medida incluyen pesos, velocidades e ingresos.

Observe que los datos al nivel nominal u ordinal de medida siempre son cualitativos, mientras que los datos al nivel de intervalo o razón siempre son cuantitativos (y por tanto pueden ser continuos o discretos). La figura 2.1 resume los posibles tipos de datos y los niveles de medida.

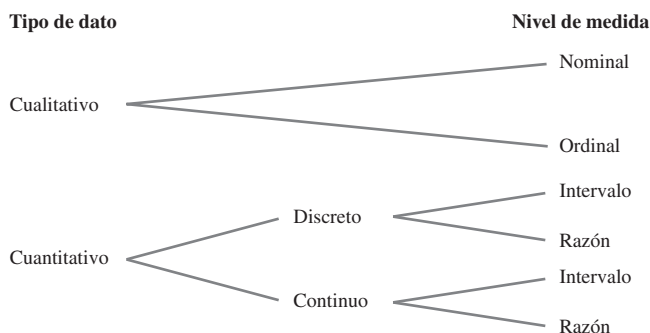


Figura 2.1 Tipos de datos y niveles de medida.

Niveles de medida

En el **nivel de medida nominal** los datos consisten sólo en nombres, etiquetas o categorías. Los datos son cualitativos y no son clasificables ni ordenables.

El **nivel de medida ordinal** aplica a datos cualitativos que se acomodan en algún orden (tal como de bajo a alto). Por lo común no tiene sentido hacer cálculos con datos en este nivel.

El **nivel de medida de intervalo** aplica a datos cuantitativos en los que los intervalos tienen significado, pero razones o cocientes no. Los datos en este nivel tienen un punto cero arbitrario.

El **nivel de medida de razón** aplica a datos cuantitativos en los que tanto los intervalos como las razones tienen sentido. Los datos en este nivel tienen un punto cero verdadero.

EJEMPLO 3 Niveles de medida

Identifique el nivel de medida (nominal, ordinal, intervalo, razón) para cada uno de los conjuntos de datos siguientes.

- Números en uniformes que identifican a los jugadores en un equipo de básquetbol.
- Clasificación que los estudiantes dan a los alimentos de la cafetería como excelente, bueno, regular o malo.
- Años del calendario para acontecimientos históricos, tal como 1776, 1945 o 2001.
- Temperatura en la escala Celsius (o centígrada).
- Tiempos de los corredores en la maratón de Boston.

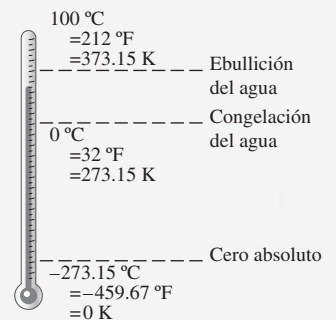
Solución

- Como se analizó en el ejemplo 1, los números de uniforme no cuentan ni miden algo. Están en el nivel de medida nominal ya que son etiquetas y no implican alguna clase de clasificación.
- Un conjunto de clasificaciones representa datos en el nivel de medida ordinal, ya que las categorías (excelente, buena, regular y mala) tienen un orden claro.
- Un intervalo en un año calendario siempre tiene el mismo significado. Pero razones o cocientes de años calendario no tienen sentido, ya que la elección del año 0 es arbitrario y no significa “el inicio del tiempo”. Por tanto, los años de calendario están en el nivel de intervalo de medida.
- Al igual que las temperaturas Fahrenheit, las temperaturas Celsius están en el nivel de medida de intervalo. Un intervalo de 1°C siempre tiene el mismo significado, pero el punto cero (0°C = punto de congelación del agua) es arbitrario y no significa “no calor”.
- La maratón tiene razones con significado, por ejemplo, un tiempo de 6 horas en realidad es el doble del tiempo que 3 horas, ya que tienen un punto cero verdadero en un tiempo de 0 horas.

Apropósito. . .

Con frecuencia los científicos miden las temperaturas en la escala Kelvin. Los datos en la escala Kelvin están en el nivel de razón de medida, ya que la escala Kelvin tiene un cero verdadero. Una temperatura de 0 kelvin en realidad es la temperatura más fría posible. Denominado *cero absoluto*, 0 K es equivalente a alrededor de -273.15°C o -459.67°F . (El símbolo de grados no se utiliza para las temperaturas Kelvin).

Escala de temperatura



Sección 2.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Datos cualitativos frente a datos cuantitativos.** ¿Cuál es la diferencia entre datos cualitativos y datos cuantitativos?
- Datos cualitativos frente a datos cuantitativos.** Un jugador de fútbol está tomando un curso de estadística y establece que los nombres de los jugadores en su equipo son cualitativos, pero pueden ser cuantitativos si se utilizan los números en las camisetas de sus uniformes. ¿Está en lo correcto? ¿Por qué sí o por qué no?
- Datos cualitativos frente a datos cuantitativos.** ¿Una investigadora está en lo correcto cuando argumenta que todos los datos son cualitativos o cuantitativos? Explique.
- Códigos postales.** Un investigador argumenta que los códigos postales son datos cuantitativos porque miden una ubicación, con números menores en el este y mayores en el oeste. ¿Está en lo correcto? ¿Por qué sí o por qué no?

Conceptos y aplicaciones

Datos cualitativos frente a datos cuantitativos. En los ejercicios del 5 al 16, determine si los datos descritos son cualitativos o cuantitativos y explique por qué.

- Grupos sanguíneos.** Los grupos sanguíneos A, B, AB y O.
- Pesos.** Los pesos de los sujetos en una clínica piloto de un nuevo medicamento.
- Estaturas.** Las estaturas de los sujetos en una clínica piloto de un nuevo medicamento.
- Películas.** Los tipos de películas (drama, comedia, etcétera).
- Duración de películas.** Las duraciones (en minutos) de películas.
- Respuestas a encuestas.** Las respuestas (sí, no) de sujetos encuestados cuando se les realizó una pregunta.
- Índices de audiencia Nielsen.** Los programas de televisión que serán observados por televidentes encuestados por Nielsen Media Research.

12. **Índices de audiencia Nielsen.** El número de hogares con televisor en uso cuando fueron encuestados por Nielsen Media Research.
13. **Rechazo a encuestas.** El número de personas que rehúsan responder preguntas cuando son encuestadas.
14. **Medidas de zapatos.** Las medidas de zapatos (tal como 8 o $10\frac{1}{2}$) de sujetos de prueba.
15. **Salario.** Los salarios de todos los gobernadores de los estados.
16. **Códigos de área.** Los códigos de área (como 617) de los teléfonos de los sujetos de una encuesta.

Datos discretos frente a datos continuos. En los ejercicios 17 a 28, establezca si los datos reales son discretos o continuos y explique por qué.

17. **Botones para peatones.** En la ciudad de Nueva York existen 3 250 botones para peatones que ellos pueden presionar en cruceros y 2 500 de estos botones no funcionan (con base en información del artículo “For Exercise in New York Futility, Push Button”, de Michael Luo, *New York Times*).
18. **Pesos de monedas.** Los pesos de las monedas de un centavo que se acuñan actualmente en Estados Unidos.
19. **Número de monedas de un centavo.** El número de monedas de un centavo recolectadas el día de hoy en todas las tiendas Sears en Estados Unidos.
20. **Encuesta Gallup.** En encuestas Gallup tomadas en diferentes instantes, sujetos que fueron cuestionados acerca de si poseían una pistola y el número de respuestas *sí* que se registraron para cada instante.
21. **Química.** Un experimento en química fue repetido y se registró el número de veces necesario para que ocurriese una reacción.
22. **Tiempos de examen.** Los tiempos necesarios para que los estudiantes terminen un examen de estadística.
23. **Calificación de exámenes.** Las calificaciones numéricas de un examen de estadística, que son el número de respuestas correctas en un examen con 25 preguntas de opción múltiple.
24. **Conteo de tránsito.** El número de automóviles que cruzan el puente Golden Gate cada hora.
25. **Velocidades de automóviles.** Las velocidades de automóviles cuando pasan por el centro del puente Golden Gate.
26. **Calificaciones a películas.** Las calificaciones a películas por un crítico: 0 estrellas, $\frac{1}{2}$ estrella, 1 estrella, etcétera.
27. **Estrellas.** El número de estrellas en cada galaxia del Universo.
28. **Calorías.** Los números de calorías consumidas por jugadores de fútbol durante un juego.
29. **Estaturas.** Las estaturas de todos los jugadores del equipo de básquetbol Lakers de Los Ángeles.
30. **Calificaciones a películas.** Las recomendaciones de un crítico de cine, dadas como “debe verse”, “buena”, “regular”, “mala”, “evítela a toda costa”.
31. **Tipos de películas.** Tipos de películas (drama, comedia, etcétera).
32. **Temperaturas.** Temperaturas del cuerpo de todos los estudiantes en una clase de estadística.
33. **Automóviles.** Clasificaciones de automóviles como subcompacto, compacto, intermedio o grande.
34. **Ensayo clínico.** Resultados de un ensayo clínico que consiste en “verdadero positivo”, “falso positivo”, “verdadero negativo”, “falso negativo”.
35. **Calificaciones.** Calificaciones finales de un curso dadas por A, B, C, D, F.
36. **Pesos.** Pesos de todos los paquetes manejados por UPS el día de hoy.
37. **NSS.** Número del seguro social.
38. **Años.** Años en los que los demócratas ganaron la elección por la presidencia de Estados Unidos.
39. **Equipaje en una aerolínea.** Pesos de todos los equipajes revisados en el vuelo 15 de cierta aerolínea.
40. **Clasificación de seguridad en automóviles.** Las clasificaciones de seguridad de la revista *Consumer Reports*, dadas por 0 = menos seguro a 3 = más seguro.

¿Las razones tienen sentido? En los ejercicios del 41 al 48, determine si el enunciado dado representa una razón (o cociente) con sentido, de modo que se aplique el nivel de medida de razón. Explique.

41. **Clasificación de película.** Una película con una clasificación de 4 estrellas es el doble de buena que una con 2 estrellas.
42. **Maratones.** Una carrera de 10 kilómetros es el doble de larga que una de 5 kilómetros.
43. **Velocidad de aviones.** Un avión que se mueve a una velocidad de 450 millas por hora viaja tres veces más rápido que otro avión que vuela a 150 millas por hora.
44. **Temperaturas.** El 6 de agosto, estuvo a 80°F en la ciudad de Nueva York, por lo que estuvo el doble de caliente que el 7 de diciembre, cuando estuvo a 40°F .
45. **Pesos de peces.** Un pez pesa 2 libras, mientras que otro pesa 4 libras, por lo que el pez más pesado pesa el doble que el otro pez.
46. **Fechado con carbono.** Por medio del fechado con carbono, se determinó que una muestra de madera es el doble de antigua que otra, puesto que la primera muestra tiene 2000 años mientras que la otra tiene 1000 años.
47. **Salario.** Un empleado tiene un salario de \$150 000, que es el doble del salario de \$75 000 de otro empleado.

Niveles de medida. Para los datos descritos en los ejercicios 29 al 40, identifique el nivel de medida como nominal, ordinal, de intervalo o de razón.

48. Calificaciones del SAT. Una persona con una calificación del SAT de 2200 está el doble de calificada para una universidad que una persona con una calificación de 1100.

Clasificación completa. En los ejercicios del 49 al 56, determine si los datos descritos son cualitativos o cuantitativos y dé sus niveles de medida. Si los datos son cuantitativos, indique si son continuos o discretos. Proporcione una breve explicación.

49. Tiempos en maratón. Tiempos finales en la maratón de la ciudad de Nueva York.

50. Orden en el maratón. Posición al finalizar (tal como primero, segundo, tercero) de los corredores de una maratón.

51. Número de identificación de empleado. Números de identificación de seis dígitos generados aleatoriamente asignados a empleados de la compañía Telektronics.

52. Tiempos de servicio de empleados. La antigüedad de cada empleado en la compañía Telektronics desde que se contrató por primera vez.

53. Años de contratación de empleados. Los años en los que los empleados fueron contratados (tal como 2003, 1998, 2007).

54. Encuesta política. En una encuesta de preferencia de votantes, los partidos políticos de los encuestados se registran como números codificados 1, 2, 3, 4 ó 5 (donde 1 = demócrata, 2 = republicano, 3 = liberal, 4 = conservador, 5 = otro).

55. Clasificación de productos. La lista de clasificación de la revista *Consumer Reports* dada por “mejor compra”, “recomendada” o “no recomendada”, para cada una de varias computadoras diferentes.

56. Control de calidad. La compañía Newport Electronics prueba cada uno de los radios que fabrica y los etiqueta como aceptable o defectuoso.

2.2 Manejo de errores

Ahora que entendió los diferentes niveles y tipos de medidas, pasamos al tema de cómo tratar los errores en la medición. Primero, consideraremos los diferentes tipos de errores que pueden ocurrir, y luego analizaremos cómo considerar los errores posibles cuando se establezcan los resultados. Observe que aunque expondremos la mayoría de este análisis en términos de medidas, se aplica igual de bien a estimaciones o proyecciones, tal como estimaciones de población o ingresos proyectados para una compañía.

Los errores son los portales del descubrimiento.

—James Joyce

Tipos de error: aleatorio y sistemático

En términos informales, los errores en la medición caen en dos categorías: errores aleatorios y errores sistemáticos. Un ejemplo ilustrará la diferencia.

Suponga que trabaja en una oficina pediátrica y utiliza una báscula digital para pesar a bebés. Si alguna vez ha trabajado con bebés, sabe que no les hace muy felices que los pongan en una báscula. Sus patateos y llanto tienden a mover la báscula, haciendo que la lectura oscile. Para el caso que se muestra en la figura 2.2a podría, posiblemente, registrar el peso del bebé como entre 14.5 y 15.0 libras. Entonces decimos que el movimiento de la báscula introduce un **error aleatorio**, ya que cualquier medición particular podría ser demasiado alta o demasiado baja.

Ahora suponga que hemos medido los pesos de los bebés a lo largo de todo el día con la báscula que se muestra en la figura 2.2b. Al final del día, observa que la báscula marca 1.2 libras cuando está vacía. En ese caso cada medida que realizó tuvo un exceso de 1.2 libras. Este tipo de error se denomina **error sistemático**, ya que es causado por un error en el *sistema* de medición, un error que afecta de manera consistente (sistemática) a todas las medidas.

A propósito...

Un error sistemático en el que las medidas en una escala (por ejemplo una balanza) difieren de manera consistente de los valores verdaderos, lo denominamos *error de calibración*. Puede probar la calibración de una balanza colocando en ella pesos conocidos, tal como 0, 5, 10 y 20 libras, y asegurándose que la balanza proporciona las lecturas esperadas.

Dos tipos de errores de medición

Los **errores aleatorios** ocurren como consecuencia de eventos aleatorios e inherentemente impredecibles en los procesos de medición.

Los **errores sistemáticos** ocurren cuando existe un problema en el sistema de medición que afecta a todas las medidas de la misma manera.

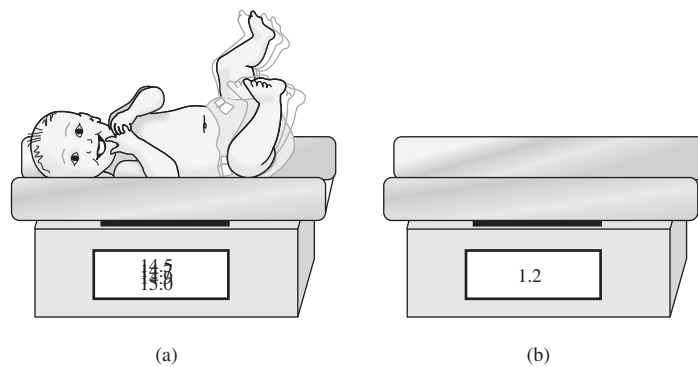


Figura 2.2 (a) El movimiento del bebé introduce errores aleatorios. (b) La báscula indica 1.2 libras cuando está vacía, con lo que introduce un error sistemático, lo cual ocasiona que todas las mediciones sean 1.2 libras más altas.

Un error sistemático afecta a todas las medidas de la misma forma, hace que todas ellas sean más altas o todas más bajas. Si descubre un error sistemático, puede regresar y ajustar las mediciones afectadas. En contraste, la naturaleza impredecible de los errores aleatorios los hace imposibles de corregir. Sin embargo, puede minimizar los efectos de errores aleatorios si hace muchas mediciones y las promedia. Por ejemplo, si mide el peso del bebé diez veces, sus mediciones tal vez serán altas en algunos casos y bajas en otros. Por tanto, puede obtener un mejor valor al promediar las diez mediciones individuales.

A propósito...

El hecho que áreas urbanas tiendan a ser más calientes de lo que sería en ausencia de actividad humana, se denomina *efecto de calor humano*. Las causas principales incluyen el calor liberado cuando los automóviles, en la casa y en la industria queman algún combustible, y por el hecho que el pavimento y los grandes edificios de concreto tienden a retener el calor de la luz solar.



UN MOMENTO DE REFLEXIÓN

Vaya a un sitio web (como www.time.gov) que proporcione la hora actual. ¿Cuán alejada está de la hora de su reloj? Describa posibles fuentes de errores aleatorio y sistemático en su control de puntualidad.

EJEMPLO 1 Errores en los datos del calentamiento global

Los científicos que estudian el calentamiento global necesitan conocer cómo la temperatura promedio de toda la Tierra, o la *temperatura promedio global*, ha cambiado con el tiempo. Considere dos dificultades al tratar de interpretar los datos de temperatura histórica desde principios del siglo XX: (1) las temperaturas se midieron con termómetros sencillos y los datos se registraron manualmente, y (2) la mayoría de las temperaturas fueron registradas en o cerca de áreas urbanas, las cuales tienden a ser mayores que las circundantes a áreas rurales a consecuencia del calor liberado por la actividad humana. Analice si cada una de estas dos dificultades produce errores aleatorios o sistemáticos, y considere las implicaciones de estos errores.

Solución La primera dificultad incluye *errores aleatorios*, ya que sin duda las personas de manera ocasional cometen errores en la lectura de los termómetros, al calibrarlos y al registrar las lecturas de las temperaturas. No hay forma de predecir si alguna lectura es correcta, más alta o más baja. Sin embargo, si hay varias lecturas para la misma región el mismo día, al promediar estas lecturas se pueden minimizar los efectos de los errores aleatorios.

La segunda dificultad incluye un *error sistemático*, ya que el exceso de calor en áreas urbanas siempre causa que la lectura de las temperaturas sea mayor de lo que debería ser de otra forma. Si los investigadores pueden estimar cuánto afectó este error sistemático a las lecturas, entonces pueden corregir los datos para este problema.

ESTUDIO DE CASO**El censo**

La constitución de Estados Unidos obliga un censo de la población cada 10 años. La Oficina del Censo de Estados Unidos realiza el censo (y también muchos otros estudios demográficos).

Al intentar contar la población, la Oficina de Censo depende en gran medida de una encuesta que es enviada y regresada por cada hogar en Estados Unidos. Sin embargo, muchos *errores aleatorios* ocurren en este proceso de encuesta. Por ejemplo, algunas personas llenan sus formas de manera incorrecta y algunas respuestas son registradas de manera errónea por los empleados de la Oficina del Censo.

El censo también está sujeto a varios tipos de *errores sistemáticos*. Por ejemplo, es muy difícil enviar encuestas a vagabundos, y es difícil contar a extranjeros indocumentados, quienes pueden ser reticentes a revelar su presencia en Estados Unidos. Estos errores sistemáticos conducen a un subconteo de la población. Un ejemplo de un error sistemático que lleva a un sobreconteo, es el conteo doble de algunos estudiantes en ciertas universidades, quienes son contados por sus padres y nuevamente en su escuela de residencia.

Por ejemplo, en el censo de 2000, originalmente se contaron 281.4 millones de personas. Sin embargo, la investigación estadística sugirió que el censo no había contabilizado a cerca de 7.6 millones de personas y contó alrededor de 4.3 millones dos veces. Si estos resultados estadísticos son correctos, entonces el censo subcontó la población por más de 3 millones de personas.

En ocasiones, una pequeña imprecisión ahorra miles de explicaciones.

—H. H. Munro (Saki)

UN MOMENTO DE REFLEXIÓN

La pregunta de si la Oficina del Censo debe permitirse ajustar su “conteo” oficial con base en encuestas estadísticas es muy polémica. La Constitución exige una “enumeración real” de la población (artículo 1, sección 2, subsección 2). ¿Usted cree que este texto impide o permite el uso de encuestas estadísticas en el conteo oficial? Defienda su opinión. También analice por qué los demócratas tienden a estar a favor del uso de métodos de muestreo mientras que los republicanos tienden a oponerse a él.

Tamaño de los errores: absoluto frente a relativo

Además de necesitar saber si un error es aleatorio o sistemático, con frecuencia queremos conocer algo acerca del tamaño total del error. Por ejemplo, ¿el error es tan grande que nos interese o es tan pequeño que no es importante? Ahora, estudiamos el concepto de tamaño del error.

Suponga que va a la tienda y compra lo que cree son 6 libras de carne para hamburguesas, pero como la balanza de la tienda está mal calibrada, en realidad sólo obtiene 4 libras. Seguramente estaría molesto por este error de 2 libras. Ahora suponga que está comprando carne para hamburguesas para barbacoa para una enorme ciudad, y ordena 3 000 libras de hamburguesas, pero en realidad recibe sólo 2 998 libras. Le faltarían las mismas 2 libras que antes, pero en este caso el error quizá no parezca muy importante. Este ejemplo sencillo muestra que el tamaño de un error depende de cómo lo valora usted.

En un lenguaje más técnico, el error de 2 libras en ambos casos es un **error absoluto**, describe cuán lejos está el valor medido del valor verdadero. Un **error relativo** compara el tamaño del error absoluto con el valor verdadero. El error relativo para el primer caso es muy grande ya que el error absoluto de 2 libras es la mitad del peso verdadero de 4 libras: por tanto, decimos que el error relativo es $2/4$ o 50%. En contraste, el error relativo del segundo caso es muy pequeño: es el error absoluto de 2 libras dividido entre el peso verdadero de la carne, que es 2 998, lo que significa un error relativo de sólo $2/2998 \approx 0.00067$ o 0.067%.

Error absoluto y error relativo

El **error absoluto** describe cuán lejos está el valor medido del valor verdadero:

$$\text{error absoluto} = \text{valor medido} - \text{valor verdadero}$$

El **error relativo** compara el tamaño del error absoluto con el valor verdadero, con frecuencia se expresa como un porcentaje:

$$\text{error relativo} = \frac{\text{error absoluto}}{\text{valor verdadero}} \times 100\%$$

Observe que los errores absoluto y relativo son *positivos* cuando el valor medido es mayor que el valor verdadero y son *negativos* cuando el valor medido es menor que el valor verdadero.

EJEMPLO 2 Error absoluto y error relativo

En cada caso determine el error absoluto y el error relativo.

- Su peso verdadero es 100 libras, pero una báscula le indica que pesa 105 libras.
- El gobierno asegura que un programa cuesta \$99 000 millones, pero una auditoría muestra que el costo verdadero es \$100 000 millones.

Solución

- El valor medido es la lectura de la báscula de 105 libras y el valor verdadero es 100 libras:

$$\begin{aligned}\text{error absoluto} &= \text{valor medido} - \text{valor verdadero} \\ &= 105 \text{ lb} - 100 \text{ lb} \\ &= 5 \text{ lb}\end{aligned}$$

$$\begin{aligned}\text{error relativo} &= \frac{\text{error absoluto}}{\text{valor verdadero}} \times 100\% \\ &= \frac{5 \text{ lb}}{100 \text{ lb}} \times 100\% \\ &= 5\%\end{aligned}$$

El peso medido se pasa por 5 libras o 5%.

Mil millones aquí, mil millones allá; dentro de poco estará hablando de dinero real.

—Senador Everett Dirksen

- El costo afirmado es \$99 000 millones y el costo verdadero es \$100 000 millones:

$$\begin{aligned}\text{error absoluto} &= \text{valor afirmado} - \text{valor verdadero} \\ &= \$99\,000 \text{ millones} - \$100\,000 \text{ millones} \\ &= -\$1\,000 \text{ millones}\end{aligned}$$

$$\begin{aligned}\text{error relativo} &= \frac{\text{error absoluto}}{\text{valor verdadero}} \times 100\% \\ &= \frac{-\$1\,000 \text{ millones}}{\$100\,000 \text{ millones}} \times 100\% \\ &= -1\%\end{aligned}$$

El costo afirmado es inferior en \$10 000 millones o 1%.

Descripción de resultados: exactitud y precisión

Una vez que se reporta una medida, debemos evaluarla para ver si es creíble a la luz de errores potenciales. En particular debemos considerar dos ideas principales acerca de cualquier valor reportado: su *exactitud* y su *precisión*. Con frecuencia, en español, estos términos se utilizan en forma intercambiable, pero en la ciencia son dos conceptos diferentes.

El objetivo de cualquier medición es obtener un valor que sea tan cercano como sea posible al *valor verdadero*. La **exactitud** describe cuán cercano es el valor medido del valor verdadero. La **precisión** describe la cantidad de detalle en la medición. Por ejemplo, suponga que un censo dice que la población de su ciudad es de 72 543, pero la población verdadera es de 96 000. El valor del censo de 72 543 es muy preciso ya que parece decimos la cantidad exacta, pero no es muy exacto ya que es casi 25% menor que la población real de 96 000. Observe que, usualmente, la exactitud es definida mediante el error relativo en lugar del error absoluto. Por ejemplo, si una compañía proyecta ventas de \$7300 millones y las ventas verdaderas resulta que son de \$7320 millones, decimos que la proyección fue muy exacta, ya que quedó a menos de 1%, aunque el error represente \$20 millones.

Definiciones

La **exactitud** describe cuánto una medida se aproxima a un valor verdadero. Una medida exacta es cercana al valor verdadero. (*Cercana* por lo general se define como un *error relativo* pequeño, en lugar de un error absoluto pequeño).

La **precisión** describe la cantidad de detalle en la *medición*.

Por lo general suponemos que la precisión con la que un número se reporta refleja cómo fue medido realmente. Si usted dice que pesa 132 libras, suponemos que midió su peso sólo a la libra más cercana. En ese caso, una medida más precisa podría determinar que usted pesó, digamos, 132.3 libras o 131.6 libras (ambas se redondearían a 132). En contraste, si dice que pesa 132.0 libras, suponemos que midió su peso al décimo de libra más cercano. Esta suposición acerca de la precisión significa que los números reportados con precisión injustificada son deshonestos. Por ejemplo, si en realidad midió su peso sólo a la libra más cercana, sería deshonesto decir que pesó 132.0 libras, ya que implicaría que hizo una medida al décimo de libra más cercano.

EJEMPLO 3 Exactitud y precisión en su peso

Suponga que su peso verdadero es de 102.4 libras. La báscula en el consultorio del doctor puede mostrar al cuarto de libra más cercano, dice que usted pesa $102\frac{1}{4}$. La báscula en el gimnasio, que proporciona lecturas digitales al 0.1 de libra más cercano, dice que usted pesa 100.7 libras. ¿Cuál báscula es más *precisa*? ¿Cuál es más *exacta*?

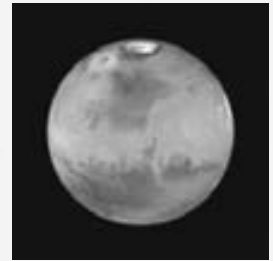
Solución La báscula en el gimnasio es más *precisa* ya que le da su peso al décimo de libra más cercano, mientras que la báscula del doctor proporciona su peso sólo al cuarto de libra más próximo. Sin embargo, la báscula en el consultorio del doctor es más *exacta* ya que su valor es más cercano a su peso verdadero.

UN MOMENTO DE REFLEXIÓN

En el ejemplo 3 necesitamos conocer su peso verdadero para determinar cuál báscula es más exacta. Pero, ¿podría conocer su peso verdadero? ¿Alguna vez puede asegurar que lo conoce? Explique.

A propósito...

En 1999 la NASA perdió el orbitador climático de Marte, de \$160 millones, cuando los ingenieros enviaron instrucciones de computadora muy precisas en unidades inglesas (libras) pero el programa de la nave las interpretó en unidades métricas (kilogramos). En otras palabras, la pérdida ocurrió porque las instrucciones muy precisas en realidad fueron muy inexactas! La NASA aprendió su lección y desde entonces ha enviado, con éxito, cuatro naves a Marte, con otro programa (denominado *Phoenix*) que descendió en Marte en 2008.



ESTUDIO DE CASO

¿El censo mide la población verdadera?

Al terminar el censo de 2000 la Oficina de Censo de Estados Unidos reportó una población de 281 421 906 (al 1 de abril de 2000), y de este modo (por la forma de reportar) implica un conteo exacto de todos los que viven en Estados Unidos. Por desgracia, tal conteo preciso quizá no sea tan exacto como parece implicar.

Incluso en principio, la única manera de obtener un conteo exacto del número de personas que viven en Estados Unidos sería contar a todos de *manera instantánea*. De otra forma, el conteo sería incorrecto ya que el número de personas cambia muy rápidamente. En Estados Unidos ocurren, en promedio, alrededor de ocho nacimientos y cuatro decesos por minuto, y un nuevo inmigrante ingresa a Estados Unidos cada tres minutos, en promedio. De hecho, la Oficina del Censo dedica muchos meses a recolectar la información que conduce al conteo. Además, los errores aleatorios y sistemáticos, con facilidad, pueden hacer impreciso el censo en unos millones de personas (vea el Estudio de caso de la página 61).

Entonces, en realidad la Oficina del Censo no podría conocer la población exacta en un día particular, y la incertidumbre en la población real es al menos de unos cuantos millones de personas. Dados estos hechos, es deshonesto reportar la población como un conteo exacto. Un reporte más honesto sería utilizar mucha menos precisión, por ejemplo, estableciendo que la población es “alrededor de 280 millones”. Para ser justos con la Oficina del Censo, sus reportes detallados explicaban las incertidumbres en el conteo de población pero estas incertidumbres rara vez se mencionaron en la prensa.

A propósito...

Los dígitos en un número que en realidad se midió se denominan *dígitos significativos*. Todos los dígitos en un número son significativos excepto por los ceros que se incluyen, por lo que el punto decimal debe colocarse de manera adecuada. Por ejemplo, 0.001234 tiene cuatro dígitos significativos, los ceros se requieren para la ubicación adecuada del punto decimal. Por la misma razón, el número 1234 000 000 tiene cuatro dígitos significativos. El número 132.0 tiene cuatro dígitos significativos; el cero es significativo ya que no es requerido para la ubicación apropiada del punto decimal y por tanto implica una medida real de cero décimos.



ESTUDIO DE CASO

Desaparecen billones en precisión injustificada de un presupuesto

A principios de 2001 los políticos anunciaron un *superávit* en el presupuesto federal de Estados Unidos de 5.6 billones de dólares durante los siguientes 10 años. En 2007, en lugar de obtener billones en dinero excedente, el gobierno había enviado otros 3 billones de dólares a la deuda. En otras palabras, con cuatro años para terminar la proyección a 10 años, la proyección ya había comprobado estar en un error por cerca de 9 billones de dólares (yendo de 5.6 billones *positivos* a 3 billones *negativos*).

Quizás usted ha escuchado o leído informes en las noticias acerca de estos muchos billones de dólares entregadas a fortunas del gobierno, y la gente de diferentes convicciones políticas pueden tener muy diferentes sentimientos con respecto a su significado. Pero aquí, para nuestros propósitos, lo más destacable fue la precisión injustificada de la proyección original. Después de todo, la proyección de \$5.6 billones fue dada al *décimo* de billón más próximo (lo cual significa a la centena de millar de millón más cercana), lo que implica que el valor real sería entre alrededor de \$5.5 y \$5.7 billones. Ahora podría decir “¡Ups!”

Para hacer justicia a los economistas responsables de la proyección, ellos estaban conscientes de las incertidumbres asociadas. Además, los cambios que ocurrieron después que se hizo la proyección —incluyendo los ataques terroristas del 11 de septiembre, la guerra en Afganistán y en Irak, y los recortes en los impuestos— hicieron comprensible que la proyección fuera tan equivocada. Pero la lección es clara. La siguiente vez que escuche a políticos o en los medios hablar de presupuestos federales futuros con gran precisión, recuerde que es muy probable que sus números resulten muy imprecisos.

Resumen: manejo de errores

Las ideas que hemos tratado en esta sección son un poco técnicas, pero muy importantes para la comprensión de medidas y errores. A continuación resumimos cómo están relacionadas las ideas.

- Los errores pueden ocurrir de muchas formas, pero por lo general los clasificamos en uno de dos tipos básicos: errores aleatorios o errores sistemáticos.
- Cualquiera que sea la fuente de un error, su tamaño puede describirse en dos formas diferentes: como un error absoluto o como un error relativo.
- Una vez que se reporta una medida, podemos evaluarla en términos de su exactitud y su precisión.

Sección 2.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Tipo de error.** Cuando un técnico registra los tiempos de reacción de un conductor en un experimento, el técnico registra de manera incorrecta uno de los tiempos reales. ¿El error es aleatorio o sistemático? Explique.
- 2. Peso estándar.** Un peso estándar definido para representar exactamente un kilogramo se guarda en el Instituto Nacional de Estándares y Tecnología. Si usted coloca este peso de un kilogramo verdadero en una balanza y ésta indica que pesa 1.002 kilogramos, ¿cuál es error absoluto y cuál el error relativo de la medición?
- 3. Medición de estaturas.** Los estudiantes en una clase de estadística miden la altura de su instructor y la registran como 174.0123668 centímetros. La altura verdadera del instructor es cercana a 176 centímetros. Describa la exactitud y la precisión de la altura registrada. Explique.
- 4. Población histórica.** Los medios reportaron que la población de Estados Unidos alcanzó los 300 millones a las 7:46 a.m. (hora del Este) del 17 de octubre de 2006. Describa la exactitud y la precisión de ese tiempo reportado.

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Especies de peces.** Existen 24 627 especies de peces en la Tierra.
- 6. Error relativo.** El error relativo que un microbiólogo comete al medir una célula debe ser menor que el error relativo que un astrónomo comete al medir una galaxia, ya que las células son más pequeñas que las galaxias.
- 7. Error de escaneo.** El administrador del supermercado Jenkins asegura que los errores de escaneo en los artículos comprados son aleatorios, y casi la mitad de los errores son a favor del supermercado.
- 8. Millaje en un automóvil.** Cuando se compra un automóvil, el millaje se reporta como 22.3655 millas por galón. Podemos tener mucha confianza en ese valor ya que es muy preciso.

Conceptos y aplicaciones

- 9. Auditor de impuestos.** Un auditor de impuestos al revisar una devolución de impuestos busca varias clases de problemas, incluyendo estas dos: (1) errores cometidos al introducir o calcular los números en la devolución de impuestos

y (2) lugares donde el contribuyente reportó ingresos de manera deshonesta. Analice si cada problema incluye errores aleatorios o sistemáticos.

- 10. Seguridad en viajes aéreos.** Antes de despegar, se supone que un piloto fija el altímetro de la aeronave a la altitud del aeropuerto. Un piloto sale de Denver (altitud 5280 pies) con su altímetro fijado en 2500 pies. Explique cómo afecta las lecturas del altímetro a lo largo de todo el vuelo. ¿Qué clase de error es éste?
 - 11. Especificaciones técnicas.** La compañía manufacturera Newport recibe un pedido por 10000 tornillos de metal, que se supone son de 25 milímetros de largo. Después que se fabrican los tornillos, un analista de control de calidad hace mediciones cuidadosas y determina que la longitud media de los tornillos es de 25 milímetros, pero alrededor de la mitad son más largos de 25 milímetros y la mitad son más pequeños que 25 milímetros. Aquí, ¿con cuál tipo de error en la medición estamos enfrentando? Explique.
 - 12. Datos de suicidas.** Con respecto a la colección de datos de suicidios, el *New York Times Almanac* asegura, “La mayoría de los expertos creen que las estadísticas de suicidios son más sombrías que las reportadas. Ellas afirman que numerosos suicidios se clasifican como accidentes y otras muertes para apiadarse de la familia”. Si éste es el caso, ¿qué clase de error se introduce en los datos de suicidios y cómo afecta los valores de los datos?
- Fuentes de error.** Para cada medición descrita en los ejercicios 13 a 20, identifique al menos una posible fuente de errores aleatorios y también identifique al menos una probable fuente de errores sistemáticos.
- 13. Encuesta.** Los ingresos anuales de 200 personas encuestadas por teléfono.
 - 14. Devolución de impuestos.** Los ingresos anuales de 200 personas son obtenidas de sus devoluciones de impuestos.
 - 15. Pesos de pasajeros.** La Administración Federal de Aviación pide a las aerolíneas pesar una muestra de pasajeros, y esos pasajeros fueron pesados con ropa y equipaje de mano.
 - 16. M&M.** Los pesos individuales de los dulces M&M resultan de colocar cada dulce en una taza de papel y luego pesar tomando en cuenta el peso de la taza.
 - 17. Radar de velocidades.** Las velocidades de automóviles las registró un oficial de policía, quien utilizó una pistola de radar.
 - 18. Productos falsificados.** El comisionado de policía en la ciudad de Nueva York estima el valor anual de los bienes falsificados en la ciudad.

19. Ventas de cigarros. La comisionada de salud de California estima el número de cigarros fumados en su estado a partir de determinar el número de cajetillas vendidas con base en las estampillas de impuestos que se ponen en cada cajetilla.

20. Medición de longitud. Un jardinero mide la longitud y el ancho de un campo de fútbol usando una regla que tiene una longitud de un pie.

Errores absoluto y relativo. En los ejercicios 21 a 24, determine los valores de los errores absoluto y relativo.

21. Factura de tarjeta de crédito. Uno de los autores recibió una factura de la tarjeta de crédito Visa por \$2995, pero incluía un cargo de \$1750 que no era válido.

22. Peso de un filete. Un estudiante de uno de los autores pesó un filete y determinó que pesaba 18 onzas, pero el menú decía que era de 20 onzas.

23. Millaje en un automóvil. La etiqueta en un automóvil establece que se obtienen 22 millas por galón en la carretera, pero las mediciones muestran que el millaje real es 24 millas por galón.

24. Pilon en la panadería. La carta de una panadería muestra que son 12 donas en una bolsa, pero el panadero por sistema pone 13 donas en cada bolsa.

25. Minimización de errores. A veinticinco personas, incluyéndolo a usted, se les pide que midan la longitud de una habitación al décimo de milímetro más cercano. Suponga que todos utilizan el mismo dispositivo de medición, que está bien calibrado, tal como una cinta métrica.

- Es muy probable que las 25 mediciones no sean exactamente iguales; así que las mediciones tendrán algunas fuentes de errores. Estos errores, ¿son aleatorios o sistemáticos? Explique.
- Si quiere minimizar el efecto de los errores aleatorios en la determinación de la longitud de la habitación, ¿qué es mejor: reportar su medición como la longitud de la habitación o reportar el promedio de las 25 mediciones? Explique.
- Describa posibles fuentes de error sistemático en la medición de la longitud de la habitación.
- ¿El proceso de promediar las 25 medidas puede ayudar a reducir los errores sistemáticos? ¿Por qué sí o por qué no?

26. Minimización de errores. El instrumento de medición para pesar la batería modelo 22F de automóvil es muy preciso y el peso se obtiene 10 veces consecutivas.

- No es probable que las diez mediciones sean exactamente iguales; así, las mediciones tendrán algunos errores. ¿Los errores son sistemáticos o aleatorios? Explique.
- Si quiere minimizar el efecto de errores aleatorios en la determinación del peso verdadero de la batería, ¿cuál es

la mejor opción: elegir una de las 10 mediciones al azar o reportar el promedio de las 10 mediciones? Explique.

c. Describa cualquier fuente de posibles errores sistemáticos en las 10 mediciones.

d. ¿El proceso de promediar los 10 pesos puede ayudar a reducir los errores sistemáticos? ¿Por qué sí o por qué no?

27. Exactitud y precisión en el peso de un Corvette. Un nuevo Corvette pesa 3273 libras. La báscula del fabricante, que es precisa a la decena de libra más próxima, proporciona el peso como 3250 libras, mientras que el Departamento de Transporte de Estados Unidos utiliza una báscula que es precisa al 0.1 de libra y obtiene un peso de 3298.2 libras. ¿Cuál medición es más *precisa*? ¿Cuál es más *exacta*? Explique.

28. Exactitud y precisión en estaturas. Suponga que el ex presidente George W. Bush tiene una estatura exacta de 71 pulgadas. Su estatura se mide con una cinta métrica que puede leerse al octavo de pulgada más próximo y el resultado que se encuentra es $71\frac{1}{8}$. Con un nuevo dispositivo láser en el consultorio médico que proporciona lecturas al 0.05 más próximo de pulgada, su altura se determinó que es 71.15 pulgadas. ¿Cuál medición es más *precisa*? ¿Cuál es más *exacta*? Explique.

29. Exactitud y precisión en pesos. Suponga que usted pesa 52.55 kilogramos. Una báscula en la clínica de salud, que proporciona mediciones de peso al medio de kilogramo más próximo, indica su peso como 53 kilogramos. Una báscula digital en el gimnasio que proporciona lecturas al 0.01 kilogramo más próximo, indica su peso como 52.88 kilogramos. ¿Cuál medición es más *precisa*? ¿Cuál es más *exacta*? Explique.

30. Exactitud y precisión en peso. Suponga que usted pesa 52.55 kilogramos. Una báscula en la clínica de salud, que proporciona mediciones de peso al medio de kilogramo más próximo, indica su peso como $52\frac{1}{2}$ kilogramos. Una báscula digital en el gimnasio que proporciona lecturas al 0.01 kilogramo más próximo, indica su peso como 51.48 kilogramos. ¿Cuál medición es más *precisa*? ¿Cuál es más *exacta*? Explique.

¿Hechos creíbles? Los ejercicios 31 a 38 proporcionan enunciados de “hechos” que provienen de mediciones estadísticas. Para cada enunciado analice brevemente fuentes posibles de error en la medición. Luego, considerando la precisión con la que se da la medición, analice si considera que el hecho es creíble.

31. Población. La población de Estados Unidos en 1860 fue de 31 443 321.

32. Muertes en vehículos de motor. El año pasado hubo 44 758 muertes en Estados Unidos debido a accidentes automovilísticos.

33. Población de China. El año pasado, la población de China fue de 1 372 557 236 personas.

- 34. Edificio más alto.** El edificio Taipei 101 tiene una altura de 1671 pies, haciéndolo el edificio más alto del mundo.
- 35.** El Puente de Entrada a San Luis tiene una altura de 630.2377599694 pies.
- 36. Teléfonos celulares.** InfoWorld reportó que actualmente el número de usuarios de teléfonos celulares es de 210 millones.
- 37. Estudiantes de universidad.** Actualmente hay 18 000 000 estudiantes en universidades en Estados Unidos.
- 38. Especies amenazadas.** De acuerdo con el gobierno de Estados Unidos ahora hay 1879 especies de animales y plantas amenazadas o en peligro de extinción.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 2 en www.aw.com/bbt.

- 39. El censo de 2010.** Vaya al sitio web de la Oficina del Censo de Estados Unidos y aprenda acerca de los planes que tuvieron para el censo de 2010. ¿Cómo y cuándo los datos fueron recolectados? ¿Existen cambios significativos en el proceso de recolección planeado (comparado con el utilizado en 2000)? En general, ¿parece más probable que el censo de 2010 será más preciso que el censo de 2000? Defienda su opinión.
- 40. Controversias en los censos.** Utilice el sitio web de la Biblioteca del Congreso de “Thomas” para determinar acerca de la legislación pendiente respecto a la recolección o uso de los datos del censo. Si encuentra más de un asunto pendiente elija uno para estudiar a profundidad. Resuma la legislación propuesta y analice brevemente tanto los argumentos a favor como los argumentos en contra.
- 41. Errores en relojes de pulsera.** Utilice un sitio web que le proporcione la hora local (tal como www.time.gov) para

fixar un reloj al segundo más cercano. Luego compare la hora de su reloj con la hora de los relojes de sus amigos. Registre los errores con signos positivos para relojes que estén adelantados de la hora verdadera y con signos negativos para los relojes que estén atrasados de la hora verdadera. Utilice los conceptos de esta sección para describir la exactitud de los relojes en su muestra.

EN LAS NOTICIAS

- 42. Errores aleatorios y errores sistemáticos.** Encuentre un artículo reciente que dé una cantidad que fue medida de manera estadística (por ejemplo, un reporte de población, ingreso promedio o número de personas sin hogar). Escriba una descripción corta de cómo fue medida la cantidad y describa brevemente cualesquiera fuentes de errores, ya sean aleatorios o sistemáticos. En general, ¿considera que la medición reportada fue exacta? ¿Por qué sí o por qué no?
- 43. Errores absolutos y errores relativos.** Encuentre un artículo reciente que describa algunos errores en una medición, estimación o número proyectado (por ejemplo, una proyección de presupuesto que salió incorrecto). En palabras, describa el tamaño del error en términos de error absoluto y error relativo.
- 44. Exactitud y precisión.** Determine un artículo reciente que le provoque duda acerca de la exactitud o precisión. Por ejemplo, el artículo podría reportar una cifra con más precisión de la que usted considera que está justificada, o podría citar una cifra que sepa que es imprecisa. Escriba un resumen de una página del artículo y explique por qué pone en duda su exactitud o precisión (o ambas).

2.3 Uso de porcentajes en estadística

Con frecuencia los resultados estadísticos se indican con porcentajes. Un porcentaje es sólo una forma de expresar una fracción; las palabras *por ciento* literalmente significan “dividido entre 100”. Sin embargo, con frecuencia los porcentajes se utilizan de manera más sutil. Considere un enunciado que apareció en un artículo de primera plana del *New York Times*:

La tasa [de fumadores] en alumnos de décimo grado saltó 45 por ciento, a 18.3 por ciento, y la tasa para alumnos de octavo grado subió a 44 por ciento, de 10.4 por ciento.

Todos los porcentajes en este enunciado son usados correctamente, pero son muy confusos. Por ejemplo, ¿qué significa decir “subió 44%, a 10.4%”? En esta sección estudiaremos algunos de los usos sutiles y abusos de los porcentajes. Antes de iniciar, debe revisar las siguientes reglas básicas para la conversión entre fracciones y porcentajes.

Conversiones entre fracciones y porcentajes

Para convertir un porcentaje a una fracción común: remplace el símbolo % con la división entre 100, si es necesario simplifique la fracción.

$$\text{Ejemplo: } 25\% = \frac{25}{100} = \frac{1}{4}$$

Para convertir un porcentaje a un decimal: quite el símbolo % y mueva el punto decimal dos lugares hacia la izquierda (es decir, divida entre 100).

$$\text{Ejemplo: } 25\% = 0.25$$

Para convertir un decimal a porcentaje: mueva el punto decimal dos lugares hacia la derecha (esto es, multiplique por 100) y agregue el símbolo %.

$$\text{Ejemplo: } 0.43 = 43\%$$

Para convertir una fracción común a un porcentaje: primero convierta la fracción común a un decimal; luego convierta el decimal a un porcentaje.

$$\text{Ejemplo: } \frac{1}{5} = 0.2 = 20\%$$

EJEMPLO 1 Encuesta en periódico

Un periódico informa que 44% de 1069 personas encuestadas dijeron que el presidente está haciendo un buen trabajo. ¿Cuántas personas dijeron que el presidente está haciendo un buen trabajo?

Solución El 44% representa la fracción de encuestados que dijeron que el presidente está haciendo un buen trabajo. Puesto que “de” por lo regular indica multiplicación, multipliquemos:

$$44\% \times 1069 = 0.44 \times 1069 = 470.36 \approx 470.$$

Alrededor de 470 de las 1069 personas dijeron que el presidente está haciendo un buen trabajo. Observe que redondeamos la respuesta a 470 para obtener un número entero de personas. (El símbolo $\approx$ significa “aproximadamente igual a”).

Uso de porcentajes para describir cambio

En estadística con frecuencia los porcentajes describen cómo cambian los datos con el tiempo. Como un ejemplo, suponga que la población de un pueblo fue 10000 en 1970 y 15000 en 2000. Podemos expresar el cambio en la población de dos formas básicas:

- Puesto que la población se elevó en 5000 personas (de 10000 a 15000), decimos que el **cambio absoluto** en la población fue de 5000 personas.
- Puesto que el aumento de 5000 personas fue 50% de la población inicial de 10000, decimos que el **cambio relativo** en la población fue 50%.

En general, al calcular un cambio absoluto o uno relativo siempre incluye dos números: un número inicial, o **valor de referencia**, y un **valor nuevo**. Una vez que identifiquemos estos dos valores, podemos calcular el cambio absoluto y el relativo con las fórmulas siguientes. Observe que los cambios absolutos y relativos son positivos si el valor nuevo es mayor que el valor de referencia y es negativo si el valor nuevo es menor que el valor de referencia.

Cambio absoluto y cambio relativo

El **cambio absoluto** describe el aumento o disminución real de un valor de referencia a un valor nuevo:

$$\text{cambio absoluto} = \text{valor nuevo} - \text{valor de referencia}$$

El **cambio relativo** describe el tamaño del cambio absoluto en comparación con el valor de referencia y puede expresarse como un porcentaje:

$$\text{cambio relativo} = \frac{\text{valor nuevo} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\%$$

UN MOMENTO DE REFLEXIÓN

Compare las fórmulas para el cambio absoluto y el relativo con las fórmulas para el error absoluto y el error relativo dadas en la sección 2.2. Describa brevemente las semejanzas que observe.

EJEMPLO 2 Crecimiento de la población mundial

La población mundial en 1950 fue 2 600 millones. Al inicio de 2000 había llegado a 6 000 millones. Describa el cambio absoluto y el relativo en la población mundial de 1950 a 2000.

Solución El valor de referencia es la población de 2 600 millones en 1950 y el valor nuevo es de 6 000 millones en 2000.

$$\begin{aligned} \text{cambio absoluto} &= \text{valor nuevo} - \text{valor de referencia} \\ &= 6\,000 \text{ millones} - 2\,600 \text{ millones} \\ &= 3\,400 \text{ millones} \\ \text{cambio relativo} &= \frac{\text{valor nuevo} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\% \\ &= \frac{6\,000 \text{ millones} - 2\,600 \text{ millones}}{2\,600 \text{ millones}} \times 100\% \\ &= 130.7\% \end{aligned}$$

La población mundial aumentó en 3 400 millones de personas o en alrededor de 131%, de 1950 a 2000.

A propósito...

De acuerdo con las Naciones Unidas y la Oficina del Censo de Estados Unidos, se calcula que la población mundial rebasó la cifra de los 6 mil millones de habitantes a finales de 1999, sólo 12 años después de alcanzar la de 5 mil millones. En 2006, la población mundial alcanzó los 6 mil quinientos millones y, probablemente, a principios de 2012 alcanzará la de 7 mil millones de personas.

Uso de porcentajes para comparaciones

Por lo común, los porcentajes también se usan para comparar dos números. En este caso, los dos números son el valor de referencia y el valor comparado.

- El **valor de referencia** es el número que estamos utilizando como la base para la comparación.
- El **valor comparado** es el otro número, que comparamos con el valor de referencia.

Entonces, podemos expresar la diferencia absoluta o la relativa entre estos dos valores con fórmulas muy similares a las de cambios absoluto y relativo. Observe que las diferencias absoluta y relativa son positivas si el valor comparado es mayor que el valor de referencia y negativa si el valor comparado es menor que el valor de referencia.

Diferencia absoluta y diferencia relativa

La **diferencia absoluta** es la diferencia entre el valor comparado y el valor de referencia:

$$\text{diferencia absoluta} = \text{valor comparado} - \text{valor de referencia}$$

La **diferencia relativa** describe el tamaño de la diferencia absoluta en comparación con el valor de referencia y es expresada como un porcentaje:

$$\text{diferencia relativa} = \frac{\text{valor comparado} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\%$$

Apropósito...

Nadie sabe todas las razones de la baja esperanza de vida de los hombres rusos, pero un factor que contribuye seguramente es el alcoholismo, que es mucho más común en Rusia que en Estados Unidos.

**EJEMPLO 3 Esperanza de vida de estadounidenses y rusos**

La esperanza de vida de los hombres estadounidenses es alrededor de 75 años, mientras que la esperanza de vida de los hombres rusos es alrededor de 59 años. Compare la esperanza de vida de los hombres estadounidenses con la de los hombres rusos en términos absolutos y relativos. (Vea la sección 6.4 para un estudio del significado de esperanza de vida).

Solución Se nos pide comparar la esperanza de vida masculina de los estadounidenses con la esperanza de vida masculina en Rusia, lo cual significa que utilizamos la esperanza de vida masculina en Rusia como el valor de referencia y la esperanza de vida masculina en Estados Unidos como el valor comparado:

$$\begin{aligned} \text{diferencia absoluta} &= \text{valor comparado} - \text{valor de referencia} \\ &= 75 \text{ años} - 59 \text{ años} \\ &= 16 \text{ años} \end{aligned}$$

$$\begin{aligned} \text{diferencia relativa} &= \frac{\text{valor comparado} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\% \\ &= \frac{75 \text{ años} - 59 \text{ años}}{59 \text{ años}} \times 100\% \\ &= 27\% \end{aligned}$$

La esperanza de vida de hombres estadounidenses es 16 años mayor, en términos absolutos, y 27% mayor, en términos relativos, que la esperanza de vida de hombres en Rusia.

De en comparación con *Más que*

Una sutileza al tratar con enunciados de porcentaje proviene de la forma como están escritos. Considere una población que *triplica* el tamaño de 200 a 600. Existen dos maneras equivalentes para indicar este cambio con porcentaje:

- Con el uso de *más que*: la población nueva es 200% *más que* la población original. Aquí estamos buscando el cambio relativo en la población:

$$\begin{aligned} \text{cambio relativo} &= \frac{\text{valor nuevo} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\% \\ &= \frac{600 - 200}{200} \times 100\% \\ &= 200\% \end{aligned}$$

- Usando *de*: la nueva población es 300% de la población original, lo que significa que es tres veces de la población original. Aquí estamos buscando la *razón* de la población nueva a la población original:

$$\frac{\text{población nueva}}{\text{población original}} = \frac{600}{200} = 3.00 = 300\%$$

Observe que los porcentajes en los enunciados “más que” y “de” están relacionados por $300\% = 100\% + 200\%$. Esto lleva a la relación general siguiente.

De en comparación con Más que (o menos que)

- Si el valor nuevo o el comparado es *P% más que* el **valor de referencia**, entonces $(100 + P)\%$ *del* valor de referencia.
- Si el valor nuevo o el comparado es *P% menos que* el valor de referencia, entonces es $(100 - P)\%$ *del* valor de referencia.

Por ejemplo, 40% *más que* el valor de referencia es 140% *del* valor de referencia, y 40% *menos que* el valor de referencia es 60% *del* valor de referencia. Cuando oiga estadísticas citadas con porcentajes, es muy importante escuchar con cuidado las palabras clave *de* y *más que* (o *menos que*) y esperemos que el orador conozca la diferencia.

EJEMPLO 4 Población mundial

En el ejemplo 2 determinamos que la población mundial en 2000 fue alrededor de 131% más que la población en 1950. Exprese este cambio con un enunciado “de”.

Solución La población mundial en 2000 fue 131% más que la población mundial en 1950. Puesto que $(100 + 131)\% = 231\%$, la población en 2000 fue 231% *de* la población en 1950. Esto significa que la población en 2000 fue 2.31 veces la población de 1950.

EJEMPLO 5 ¡Oferta!

Un almacén tiene una venta con “25% de descuento”. ¿Cómo se compara un precio de oferta con un precio original?

Solución El “25% de descuento” significa que el precio de venta es 25% *menos que* el precio original, lo cual significa que es $(100 - 25)\% = 75\%$ *del* precio original. Por ejemplo, si el precio original de un artículo fue \$100, su precio de venta es \$75.

UN MOMENTO DE REFLEXIÓN

Una tienda anuncia “¡1/3 de descuento en toda la tienda!”. Otra tienda anuncia “¡Precios de venta a sólo 1/3 de los precios originales!” ¿Cuál tienda tiene precios más rebajados? Explique.

Porcentajes de porcentajes

Los cambios de porcentaje y las diferencias de porcentajes pueden ser muy confusos cuando los valores *mismos* son porcentajes. Suponga que su banco aumenta la tasa de interés en su cuenta de ahorro de 3 a 4%. Uno se siente tentado a decir que la tasa de interés aumentó en 1%,

Apropósito. . .

En la actualidad la población mundial está creciendo en alrededor de 75 millones de personas cada año, lo que significa que en sólo cuatro años el mundo aumentará aproximadamente en tanta gente como toda la población de Estados Unidos.



pero en el mejor de los casos el enunciado es ambiguo. La tasa de interés aumentó en 1 *punto porcentual*, pero el cambio relativo en la tasa de interés es 33%:

$$\frac{4\% - 3\%}{3\%} \times 100\% = 0.33 \times 100\% = 33\%$$

Por tanto, puede decir que el banco elevó su tasa de interés en 33%, aunque la tasa real aumentó en sólo 1 punto porcentual (de 3 a 4%).

Puntos porcentuales en comparación con %

Cuando ve un cambio o diferencia expresado en *puntos porcentuales*, puede suponer que es un cambio o diferencia *absoluto*. Si se expresa como un porcentaje, probablemente es un cambio o diferencia *relativo*.

EJEMPLO 6 Margen de error

Con base en entrevistas con una muestra de estudiantes en su escuela, usted concluye que el porcentaje de todos los estudiantes que son vegetarianos quizás esté entre 20 y 30%. ¿Debe reportar su resultado como “25% con un margen de error de 5%” o como “25% con un margen de error de 5 puntos porcentuales”? Explique. (Vea la sección 1.1 para revisar el significado de margen de error).

Solución El rango de 20 a 30% proviene de restar y sumar una *diferencia absoluta* de 5 puntos porcentuales a 25%. Esto es,

$$20\% = (25 - 5)\% \quad \text{y} \quad 30\% = (25 + 5)\%$$

Por tanto, el enunciado correcto es “25% con un margen de error de 5 puntos porcentuales”. Si en su lugar usted dice “25% con un margen de error de 5%” usted deduciría que el error era 5% de 25%, que sólo es 1.25%.

*Si no puede convencerlos,
confúndalos.*

—Harry S. Truman,
ex presidente de Estados Unidos

EJEMPLO 7 Cuidado en la redacción

Suponga que 40% de los votantes registrados en la ciudad de Carson son republicanos. Lea cuidadosamente las preguntas siguientes y proporcione las respuestas más apropiadas.

- El porcentaje de votantes registrados como republicanos es 25% mayor en Freetown que en Carson City. ¿Qué porcentaje de los votantes registrados en Freetown son republicanos?
- El porcentaje de votantes registrados como republicanos es 25 puntos porcentuales mayor en Freetown que en Carson City. ¿Qué porcentaje de los votantes registrados en Freetown son republicanos?

Solución

- Interpretamos “25%” como una diferencia relativa, y 25% de 40% es 10% (ya que $0.25 \times 0.40 = 0.10$). Por tanto, el porcentaje de republicanos registrados en Freetown es $40\% + 10\% = 50\%$.
- En este caso, interpretamos los “25 puntos porcentuales” como una diferencia absoluta, por lo que sólo sumamos este valor al porcentaje de republicanos en Carson City. Por tanto, el porcentaje de republicanos registrados en Freetown es $40\% + 25\% = 65\%$.

Sección 2.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Interpretación de porcentaje.** Tenga en cuenta la cita del comienzo de esta sección: “la tasa (de fumar) entre los estudiantes universitarios pasó de 45 a 18.3%, y la tasa de estudiantes de preparatoria pasó del 44 a 10.4%”. Explique brevemente el significado de cada uno de los porcentajes en este comunicado.
- Porcentaje.** Un editorial del *New York Times* criticó un comercial que describe un enjuague bucal como algo que “reduce la placa en los dientes por más de 300%”. Si el enjuague dental elimina toda la placa, ¿qué porcentaje se remueve? ¿Es posible reducir la placa en más de 300%?
- Puntos porcentuales.** Una encuesta de Gallup, donde participan 1236 adultos, nos señala que 5% de ellos cree que se tiene mala suerte si se rompe un espejo. Este porcentaje tiene un índice de error de 1.2 puntos porcentuales. ¿Por qué es erróneo afirmar que el porcentaje es de 5% con un margen de error de 1.2%?
- De y más que.** Una encuesta incluye 1400 encuestados. ¿Cuánto es 5% de 1400? ¿Cuánto es 5% de más de 1400?

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Teléfono celular.** El porcentaje de gente con teléfono celular aumentó en 1.2 millones de personas.
- Porcentajes.** El director general de Telektronics tiene un salario anual que es 50% mayor al del director financiero, por lo que el salario del director financiero es 50% menor que el del director general.
- Tasa de interés.** El banco Jefferson Valley aumentó su tasa para préstamo de automóviles nuevos en 100%.
- Tasa de interés.** El banco Jefferson Valley aumentó su tasa (anual) para préstamos de automóviles nuevos en 100 puntos porcentuales.

Conceptos y aplicaciones

- Fracciones, decimales y porcentajes.** Exprese cada uno de los números siguientes en las tres formas: fracción, decimal y porcentaje.
 - 2/5
 - 1.50
 - 0.25
 - 30%
- Fracciones, decimales y porcentajes.** Exprese cada uno de los números siguientes en las tres formas: fracción, decimal y porcentaje.
 - 225%
 - 0.375
 - 1/20
 - 0.12

11. Práctica con porcentajes. Al realizar un estudio de las declaraciones hechas por 1028 criminales, 956 de ellos se declararon culpables y 392 de ellos fueron sentenciados a prisión. Entre los 72 criminales restantes, quienes se declararon inocentes, 58 de ellos fueron enviados a prisión (con base en información de “Does It Pay to Plead Guilty?” de Brereton y Casper, *Law and Society Review*, vol. 16, número 1).

- ¿Qué porcentaje de los criminales se declararon culpables?
- ¿Qué porcentaje de los criminales fueron enviados a prisión?
- Entre los que se declararon culpables, ¿cuál es el porcentaje que fue enviado a prisión?
- Entre los que se declararon inocentes, ¿cuál es el porcentaje que fue enviado a prisión?

12. Práctica con porcentajes. Se llevó a cabo un estudio para determinar si lanzar o girar una moneda de un centavo tiene algún efecto sobre la proporción de caras. En 49437 ensayos, 29015 fueron lanzamientos de la moneda y 14709 de esas monedas salió cara. Los otros 20422 ensayos incluyeron girar la moneda y 9197 de ellos salieron en cara (con base en datos de Robin Lock como aparece en *Chance News*).

- ¿Qué porcentaje de los ensayos incluyeron lanzamiento de monedas?
- ¿Qué porcentaje de los ensayos incluyeron hacer girar las monedas?
- Entre las monedas que fueron lanzadas, ¿cuál es el porcentaje que salió cara?
- Entre las monedas que fueron giradas, ¿cuál es el porcentaje que salió cara?

Cambio relativo. Cada uno de los ejercicios 13 a 20 proporciona dos valores. Para cada par de valores utilice un porcentaje para expresar el cambio o la diferencia relativos. Utilice el segundo valor dado como el valor de referencia y exprese los resultados al punto porcentual más próximo. Además, escriba un enunciado que describa su resultado.

- Periódico.** En la actualidad el número de periódicos en Estados Unidos es de 1456, y en 1900 fue de 226.
- Automóviles.** Ahora existen 136651000 automóviles particulares registrados y en 1980 fueron 121601000.
- Aviones.** Ahora existen 8206 aviones comerciales en Estados Unidos y en 1980 fueron 3808.
- Quiebras.** El año pasado hubo 1725300 quiebras registradas y en el año 2000 se registraron 1276900 casos de quiebras.
- Periódicos.** La circulación diaria del *Wall Street Journal* es alrededor de 1.75 millones (la mayor en Estados Unidos). La circulación diaria del *New York Times* es alrededor de 1.09 millones (la tercera mayor en Estados Unidos).

- 18. Venta de automóviles.** En un año reciente se vendieron 13 525 automóviles Chrysler 300C y 45 782 Honda híbridos.
- 19. Aeropuertos.** El aeropuerto O'Hare de Chicago tuvo un tráfico de 34 millones de pasajeros el año pasado. El aeropuerto con mayor tránsito en el mundo, el aeropuerto Hartsfield de Atlanta, dio servicio a 42 millones de pasajeros el año pasado.
- 20. Turistas.** En un año reciente Francia estaba clasificada como el destino número uno de los turistas, con 75.1 millones de llegadas internacionales. Estados Unidos estaba en tercero, con 46.1 millones de llegadas internacionales.

Encuestas. Algunos análisis de resultados de encuestas requieren que usted sepa el número real de sujetos cuyas respuestas caen en una categoría particular. En los ejercicios 21 a 24, determine el número de encuestados que corresponden al porcentaje dado.

- 21. Llamadas personales.** En una encuesta de 1385 oficinistas, 4.8% dijo que no realizaba llamadas personales.
- 22. Errores en entrevistas.** En una encuesta de 150 ejecutivos, 47% dijo que la mayoría de los errores en entrevistas comunes es tener poco o nulo conocimiento de la compañía.
- 23. Televisores.** En una encuesta de Frank N. Magid Associates a 1005 adultos, 83% reportó que tenían más de un televisor en casa.
- 24. Teléfonos celulares.** En una encuesta de Frank N. Magid Associates a 1109 consumidores mayores a 11 años de edad, 81% reportó tener al menos un teléfono celular en su hogar.

De en comparación con más que. En los ejercicios del 25 al 28, rellene los espacios en blanco. En cada caso explique brevemente su razonamiento.

- 25. Pesos.** Si un camión pesa 40% más que un automóvil, entonces el peso del camión es _____% del peso del automóvil.
- 26. Áreas.** Si el área de Noruega es 24% mayor que el área de Colorado, entonces el área de Noruega es _____% del área de Colorado.
- 27. Población.** Si la población de Montana es 20% menor que la población de New Hampshire, entonces la población de Montana es _____% de la población de New Hampshire.
- 28. Salario.** El salario actual de Jay Leno es 18% menor que el salario de David Letterman, así que el salario de Jay Leno es _____% del salario de David Letterman.
- 29. Margen de error.** Una encuesta de Gallup a 1012 adultos mostró que 89% de los estadounidenses dicen que la clonación humana no debe permitirse. El margen de error fue de 3 puntos porcentuales. ¿Importaría si un periódico reportó el margen de error como 3%? Explique.
- 30. Margen de error.** Una encuesta del Centro de Investigación Pew a 3002 adultos mostró que el porcentaje que escuchaba National Public Radio probablemente está entre 14 y 18%. ¿Cómo debe reportar el margen de error un periódico? Explique.

Porcentajes de porcentajes. Los ejercicios 31 a 34 describen cambios en los cuales las medidas mismas son porcentajes. Exprese cada cambio de dos formas (1) como una diferencia absoluta en términos de puntos porcentuales, y (2) como una diferencia relativa en términos de por ciento.

- 31.** El porcentaje de alumnos de último año en preparatoria que ingieren alcohol disminuyó de 68.2% en 1975 a 52.7% ahora.
- 32.** El porcentaje de población mundial que vive en países en desarrollo disminuyó de 27.1% en 1970 a 19.5% ahora.
- 33.** La tasa de supervivencia de cinco años para caucásicos para todas las formas de cáncer aumentó de 39% en la década de 1960 a 61% en la actualidad.
- 34.** La tasa de supervivencia de cinco años para afroamericanos para todas las formas de cáncer aumentó de 27% en la década de 1960 a 48% en la actualidad.



Proyectos para internet y más allá

Para enlaces útiles seleccione "Links for Internet Projects" para el capítulo 2 en www.aw.com/bbt.

- 35. Población mundial.** Determine la estimación actual de la población mundial en el reloj de población mundial de la Oficina del Censo de Estados Unidos. Describa el cambio porcentual en la población desde que la marca de 6 mil millones se pasó durante 1999. Además, determine cómo la estimación del reloj de población se hizo y analice las incertidumbres en la estimación de la población mundial.
- 36. Estadísticas de uso de sustancias.** Vaya al sitio web para el Centro Nacional sobre la Adicción y Abuso de Sustancias (CASA, por sus siglas en inglés) y encuentre un reporte reciente que proporcione estadísticas sobre abuso de sustancias. Escriba un resumen de una página de la nueva investigación, dando al menos alguna de las conclusiones en términos de porcentaje.

EN LAS NOTICIAS

- 37. Porcentajes.** Encuentre tres artículos recientes en los que se utilicen porcentajes para describir resultados estadísticos. En cada caso describa el significado del porcentaje.
- 38. Cambio porcentual.** Encuentre un artículo reciente en el que se utilicen porcentajes para expresar el cambio en un resultado estadístico de un instante a otro (tal como un aumento en la población o en el número de niños que fuman). Describa el significado del cambio. Asegúrese de observar las palabras clave tal como *de* o *más que*.

2.4 Números índice

Si usted escucha el reporte económico nocturno, quizá haya escuchado acerca de **números índices**, tal como el índice de precios al consumidor, el índice de precios al productor o el índice de confianza del consumidor. Los números índice son muy comunes en estadística, ya que proporcionan una manera sencilla de comparar medidas realizadas en diferentes tiempos o en distintos lugares. En esta sección investigaremos el significado y uso de los números índice, centrándonos en el índice de precios al consumidor (IPC). Iniciamos con un ejemplo que utiliza precios de gasolina.

La tabla 2.1 muestra el precio promedio de la gasolina en Estados Unidos para años seleccionados de 1955 a 2005. (Son precios reales de esos años; es decir, *no* se han ajustado a la inflación). Suponga que, en lugar de los precios mismos, queremos conocer cómo se compara el precio de la gasolina en años diferentes con el precio de 1975. Una manera de comparar los precios sería expresar el precio de cada año, como un porcentaje del precio de 1975. Por ejemplo, al dividir el precio de 1965 entre el precio de 1975, encontramos que el precio de 1965 fue 55.0% del precio de 1975:

$$\frac{\text{precio de 1965}}{\text{precio de 1975}} = \frac{31.2¢}{56.7¢} = 0.550 = 55.0\%$$

Al proceder de manera similar para cada uno de los otros años, podemos calcular todos los precios como porcentajes del precio de 1975. La tercera columna de la tabla 2.1 muestra los resultados. Observe que el porcentaje para 1975 es 100%, ya que elegimos el precio de 1975 como el valor de referencia.

Tabla 2.1 Precio promedio de gasolina (por galón)			
Año	Precio	Precio como porcentaje del precio de 1975	Precio índice (1975 = 100)
1955	29.1¢	51.3%	51.3
1965	31.2¢	55.0%	55.0
1975	56.7¢	100.0%	100.0
1985	119.6¢	210.9%	210.9
1995	120.5¢	212.5%	212.5
2000	155.0¢	273.4%	273.4
2005	231.0¢	407.4%	407.4

Fuente: Departamento de Energía de Estados Unidos.

Ahora vea la última columna de la tabla 2.1. Es idéntica a la tercera columna, salvo que hemos quitado los signos %. Este cambio sencillo convierte a los números de porcentajes a un *precio índice*, que es un tipo de número índice. El enunciado “1975 = 100” en el encabezado de la columna muestra que el valor de referencia es el precio de 1975. En este caso, en realidad no existe diferencia entre indicar las comparaciones como porcentajes y como números índice, es cuestión de elección y conveniencia. Sin embargo, como veremos dentro de poco, es tradicional utilizar números índice en lugar de porcentajes en casos donde muchos factores se consideran de manera simultánea.

A propósito...

Los precios en la tabla 2.1 son promedio de todo un año que no muestran cuánto variaron los precios. Por ejemplo, durante 2005 los precios de la gasolina tuvieron un pico muy breve, de \$3.13 por galón, a consecuencia del huracán Katrina.



Números índice

Un **número índice** proporciona una manera sencilla de comparar mediciones realizadas en tiempos diferentes o en lugares diferentes. El valor en un tiempo particular (o lugar) debe elegirse como el *valor de referencia* (o *valor base*). El número índice para cualquier otro tiempo (o lugar) es

$$\text{número índice} = \frac{\text{valor}}{\text{valor de referencia}} \times 100$$

Apropósito...

El término *índice* por lo común se utiliza para casi cualquier clase de números que proporciona una comparación útil, aunque los números no son números índice estándar. Por ejemplo, el índice de masa corporal (IMC) proporciona una manera de comparar gente por su peso y su estatura, pero está definido sin ningún valor de referencia. En específico, el índice de masa corporal se define como el peso (en kilogramos) dividido entre el cuadrado de la estatura (en metros).

Un estudio de economía por lo regular muestra que el mejor momento para comprar cualquier cosa es el año pasado.

—Marty Allen

EJEMPLO 1 Determinación de un número índice

Suponga que el costo de la gasolina el día de hoy es \$3.20 por galón. Con el precio de 1975 como valor de referencia, determine el número índice para la gasolina hoy.

Solución La tabla 2.1 muestra que el precio de la gasolina fue 56.7 centavos o \$0.567 por galón en 1975. Si utilizamos el precio de 1975 como el valor de referencia y el precio actual es \$3.20, el número índice para la gasolina hoy es

$$\text{número índice} = \frac{\text{precio actual}}{\text{precio de 1975}} \times 100 = \frac{\$3.20}{\$0.567} \times 100 = 564.4$$

Este número índice para el precio actual es 564.4. Esto significa que el precio de la gasolina actual es \$564.4% del precio de 1975.

UN MOMENTO DE REFLEXIÓN

Determine el precio real de la gasolina hoy en una gasolinera cercana. ¿Cuál es el precio índice para el precio actual con el precio de 1975 como el valor de referencia?

Comparaciones con números índice

El propósito principal de los números índice es facilitar las comparaciones. Por ejemplo, suponga que queremos saber cuánto fue más cara la gasolina en 2000 que en 1975. Con facilidad podemos obtener la respuesta de la tabla 2.1, que utiliza el precio de 1975 como el valor de referencia. Esta tabla muestra que el precio índice para 2000 fue 273.4, lo que significa que el precio de la gasolina en 2000 fue 273.4% del precio de 1975. De manera equivalente, podemos decir que el precio de 2000 fue 2.734 veces el precio de 1975.

También podemos hacer comparaciones cuando no existe ningún valor de referencia. Por ejemplo, suponga que queremos conocer cuánto es más cara la gasolina en 1995 que en 1965. Determinamos la respuesta dividiendo los números índice para los dos años:

$$\frac{\text{número índice para 1995}}{\text{número índice para 1965}} = \frac{212.5}{55.0} = 3.86$$

El precio de 1995 fue 3.86 veces el precio de 1965, o 386% del precio de 1965. En otras palabras, la misma cantidad de gasolina que cuesta \$1.00 en 1965 tendría un costo de \$3.86 en 1995.

EJEMPLO 2 Uso del índice de precio de gasolina

Utilice la tabla 2.1 para responder las preguntas siguientes.

- Suponga que en 1975 costó \$7.00 llenar su tanque de gasolina. ¿Cuánto costaría comprar la misma cantidad de gasolina en 2005?
- Suponga que en 1995 costó \$20.00 llenar su tanque de gasolina. ¿Cuánto costaría comprar la misma cantidad de gasolina en 1955?

Solución

- La tabla 2.1 muestra que el precio índice (1975 = 100) para 2005 fue 407.4, lo cual significa que el precio de la gasolina en 2005 fue 407.4% del precio de 1975. Así que el precio de gasolina que costó \$7.00 en 1975 fue

$$407.4\% \times \$7.00 = 4.074 \times \$7.00 = \$28.52$$

- b. La tabla 2.1 muestra que el precio índice (1975 = 100) para 1995 fue 212.5 y el índice para 1955 fue 51.3. Así el costo de la gasolina en 1955 comparada con el costo en 1995 fue

$$\frac{\text{número índice para 1955}}{\text{número índice para 1995}} = \frac{51.3}{212.5} = 0.2414$$

La gasolina que cuesta \$20.00 en 1995 cuesta $0.2414 \times \$20.00 \approx \4.83 en 1955.

El índice de precios al consumidor

Hemos visto que el precio de la gasolina se elevó de manera sustancial con el tiempo. La mayoría de otros precios y salarios también se elevaron, un fenómeno que llamamos **inflación**. (Los precios y salarios de manera ocasional disminuyen con el tiempo, lo cual es *deflación*). Por tanto, los cambios en el precio real de la gasolina no son muy significativos a menos que los comparemos con la tasa global de inflación, que se mide mediante el **índice de precios al consumidor (IPC)**.

El índice de precios al consumidor se calcula y reporta mensualmente por la Oficina de Estadísticas del Trabajo de Estados Unidos. Representa un promedio de precios para una muestra de bienes, servicios y vivienda. La muestra mensual consiste en más de 60 000 artículos. Los detalles de la recolección de datos y el cálculo del índice son bastante complejos, pero el IPC mismo es un número índice simple. La tabla 2.2 muestra el IPC anual promedio durante un periodo de 30 años. En nuestros días el valor de referencia para el IPC es un promedio de precios durante el periodo 1982-1984, por lo cual la tabla dice “1982-1984 = 100”.

El índice de precios al consumidor

El **índice de precios al consumidor (IPC)**, que se calcula y reporta mensualmente, tiene como base una muestra del costo de más de 60 000 bienes, servicios y vivienda.

Tabla 2.2 Índice de precios al consumidor anual promedio (1982-1984 = 100)

Año	IPC	Año	IPC	Año	IPC
1977	60.6	1987	113.6	1997	160.5
1978	65.2	1988	118.3	1998	163.0
1979	72.6	1989	124.0	1999	166.6
1980	82.4	1990	130.7	2000	172.2
1981	90.9	1991	136.2	2001	177.1
1982	96.5	1992	140.3	2002	179.9
1983	99.6	1993	144.5	2003	184.0
1984	103.9	1994	148.2	2004	188.9
1985	107.6	1995	152.4	2005	195.3
1986	109.6	1996	156.9	2006	201.6

NOTA TÉCNICA

El gobierno mide dos índices de precios al consumidor. El IPC-U con base en productos que reflejan los hábitos de compra de todos los consumidores urbanos, mientras que el IPC-W tiene como base los hábitos de compra de únicamente los asalariados. (La tabla 2.2 muestra el IPC-U).

El IPC nos permite comparar precios en tiempos diferentes. Por ejemplo, para determinar cuánto fueron mayores los precios en 2005 que en 1995, dividimos el IPC para los dos años:

$$\frac{\text{IPC para 2005}}{\text{IPC para 1995}} = \frac{195.3}{152.4} = 1.28$$

Con base en el IPC, los precios típicos en 2005 fueron 1.28 veces los de 1995. Por ejemplo, un artículo común que en 1995 costó \$1000, en promedio, tendría un costo \$1280 en 2005. Por supuesto, los artículos individuales pueden tener cambios en el precio que son diferentes del promedio. Por ejemplo, los precios de las computadoras para potencia de cómputo equivalente *cayeron* significativamente de 1995 a 2005, lo que significa que podría comprar una computadora mucho más poderosa en 2005 por el mismo o menos dinero. En contraste, el precio promedio del seguro de salud fue más del doble en el mismo periodo, por lo que decimos que los precios del seguro social se elevaron mucho más rápido que la tasa de inflación general.

EJEMPLO 3 Cambios en el IPC

Suponga que necesitó \$30 000 para mantener un estándar de vida particular en 2000. ¿Cuánto habría necesitado en 2006 para mantener el mismo estándar de vida? Suponga que el precio promedio de sus compras comunes se elevó a la misma tasa que el índice de precios al consumidor (IPC).

Solución Comparamos los IPC para 2006 y 2000:

$$\frac{\text{IPC para 2006}}{\text{IPC para 2000}} = \frac{201.6}{172.2} = 1.17$$

Esto es, los precios comunes en 2006 fueron alrededor de 1.17 veces los de 2000. Así, si usted necesitó \$30 000 en 2000, habría necesitado $1.17 \times \$30\,000 = \$35\,100$ para tener el mismo estándar de vida en 2006.

Apropósito...

Los salarios de atletas profesionales alguna vez se mantuvieron bajos ya que a los jugadores no se les permitía ofrecer sus habilidades en el mercado libre ("agencia libre"). Eso cambió después que el jugador estrella del béisbol Curt Flood presentó una demanda contra la Liga Mayor de Béisbol en 1970. En última instancia, Flood perdió cuando la Suprema Corte falló en favor del béisbol en 1972, pero el proceso que él puso en movimiento (hacia la agencia libre) fue imparable.



Ajuste de precios por la inflación

A mediados de 2000 los precios de la gasolina subieron de manera repentina, alcanzando un pico de alrededor de \$1.87 por galón. Aunque podría parecer barato para los estándares de hoy, en ese momento causó indignación entre los consumidores cuando las noticias los pregonaron como "los precios más altos de gasolina en la historia".

En términos de precios reales (como se indican en la bomba) el precio de \$1.87 por galón hace añicos el récord alto anterior de \$1.47 de 1981. Pero, ¿los precios de la gasolina *realmente* fueron un récord alto? Si queremos comparar los precios de manera justa, debemos tomar en cuenta los efectos de la inflación. Para hacer eso, primero debemos saber cómo cambiaron los precios comunes entre 1981 y 2000, lo cual podemos hacer dividiendo los IPC para esos años, usando datos de la tabla 2.2:

$$\frac{\text{IPC para 2000}}{\text{IPC para 1981}} = \frac{172.2}{90.9} = 1.89$$

Puesto que el IPC se elevó en un factor de 1.89, el precio de la gasolina en 1981 de \$1.47 fue equivalente a un precio de $\$1.47 \times 1.89 = \2.78 . En el lenguaje de economía, decimos que \$1.47 en "dólares de 1981" fue equivalente a \$2.78 en "dólares de 2000". Puesto que el precio real de la gasolina en 2000 fue mucho menor que \$2.78, el precio de 2000 *no* fue tan alto como el precio de 1981 en "términos reales", significando precios ajustados por la inflación. La figura 2.3 muestra más de 50 años de datos de precios anuales para precios reales de la gasolina y los precios ajustados a dólares de 2006. Observe que, en términos reales, el precio promedio anual de 1981 permaneció como el récord hasta 2006.

EJEMPLO 4 Salarios en el béisbol

En 1987 el salario medio para los jugadores en la Liga Mayor de Béisbol fue \$412 000. En 2006, fue \$2 867 000. Compare el incremento en los salarios medio de béisbol a la tasa general de inflación medida por el IPC.

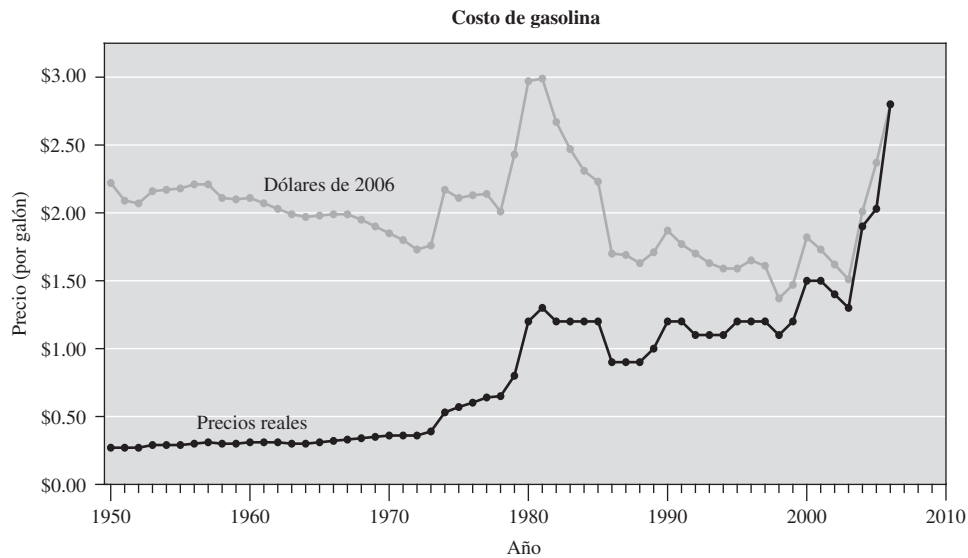


Figura 2.3 Precios de gasolina (promedio anual), 1950-2006. Observe que, como utilizamos dólares de 2006 para los precios ajustados por la inflación, los precios reales y ajustados son los mismos para 2006.
Fuente: American Petroleum Institute.

Solución Primero comparamos los índices de precios al consumidor para 2006 y 1987:

$$\frac{\text{IPC para 2006}}{\text{IPC para 1987}} = \frac{201.6}{113.6} = 1.77$$

Ahora comparamos los salarios promedio para esos dos años:

$$\frac{\text{salario promedio de béisbol para 2006}}{\text{salario promedio de béisbol para 1987}} = \frac{\$2867000}{\$412000} = 6.96$$

Durante el mismo periodo que los precios promedio (medidos mediante el IPC) se elevaron alrededor de 77%, el salario medio de béisbol se elevó casi 600%. En otras palabras, los salarios de los jugadores de la Liga Mayor de Béisbol se elevaron casi $600/77 \approx 8$ veces la tasa global de la inflación.

Otros números índice

El índice de precios al consumidor es sólo uno de muchos números índice que verá en los reportes de noticias. Algunos también son índices de precios, tal como el índice de precios al productor (IPP), el cual mide los precios que los productores (fabricantes) pagan por los bienes que compran (en lugar de los precios que el consumidor paga). Otros índices intentan medir variables más cualitativas. Por ejemplo, el índice de confianza del consumidor tiene como base una encuesta diseñada para medir actitudes de los consumidores de modo que las empresas puedan calcular si la gente gasta o ahorra. Nuevos índices se crean con frecuencia por grupos tratando de proporcionar comparaciones simples.

A propósito...

¿Piensa convertirse en comediante? Entonces probablemente necesitará verificar el índice de costo de la risa (¡En realidad existe uno!), que sigue la trayectoria de los costos de artículos tales como pollos de hule, anteojos de Groucho Marx y la admisión a clubes de humorismo.

Sección 2.4 Ejercicios

Alfabetización estadística y pensamiento crítico

- Número índice.** Un periódico reporta que el índice del precio del gas en 2007 fue \$2.27 por galón. ¿Qué es incorrecto en ese enunciado?
- Número índice.** Si los costos de una computadora en el año 2000 se fijan iguales a 100 de modo que puedan utilizarse como la base para la determinación de los números índice, y el número índice para el año 2005 es 20, ¿qué sabemos acerca de los costos de una computadora en 2005 comparados con los costos de una computadora en 2000?
- IPC.** Si los precios de los bienes, servicios y vivienda aumentan, ¿el IPC debe aumentar? Explique.
- IPC.** Si el IPC aumenta, ¿los salarios también deben aumentar? Explique.

Conceptos y aplicaciones

Índice de precios de gasolina. En los ejercicios del 5 al 8, utilice el índice de precios de gasolina de la tabla 2.1. Explique brevemente su razonamiento en cada caso.

- Datos actuales.** Suponga que el costo de la gasolina hoy es \$3.25 por galón. ¿Cuál es el número índice para el precio de la gasolina hoy, tomando el precio de 1975 como el valor de referencia?
- Índice 2006.** El precio promedio de un galón de gasolina en 2006 fue \$2.62. ¿Cuál es el número índice precio para la gasolina en 2006, con el precio de 1975 como el valor de referencia?
- Precio de 1998.** Tomando el precio de 1975 como el valor de referencia, el número índice precio de la gasolina para 1998 es 197.5. ¿Cuál fue el costo de un galón de gasolina en 1998?
- Precio 1978.** Con el precio de 1975 como el valor de referencia, el índice de precios de la gasolina para 1978 es 114.6. ¿Cuál fue el costo de un galón de gasolina en 1978?
- Reconstrucción del índice de precios de gasolina.** Identifique los siete números índice de los precios en la tabla 2.1 que resultarían de usar el precio de 1965 como el valor de referencia. (Sugerencia: cree una columna para el precio como un porcentaje del precio de 1965 y otra columna que proporcione el precio índice con 1965 = 100).
- Reconstrucción del índice de precios de gasolina.** Identifique los siete números índice de los precios en la tabla 2.1 que resultarían de usar el precio de 2000 como el valor de referencia. (Sugerencia: cree una columna para el precio como un porcentaje del precio de 2000 y otra columna que proporcione el precio índice con 2000 = 100).

- Uso del índice de precios de gasolina.** Si costó \$16.74 llenar su tanque de gasolina en 1985, ¿cuánto sería el costo de llenar el mismo tanque en 2000?
- Uso del índice de precios de gasolina.** Si costó \$23.25 llenar su tanque de gasolina en 2000, ¿cuánto sería el costo de llenar el mismo tanque en 2005?

Índice de precios al consumidor. En los ejercicios 13 al 16, utilice el índice de precios al consumidor de la tabla 2.2.

- Costo de universidades privadas.** El costo anual promedio (matrícula, cuotas, alojamiento y comida) en una universidad privada para cuatro años se elevó de \$5 900 en 1980 a \$27 516 en 2004. Calcule el porcentaje de aumento en el costo de 1980 a 2004 y compárelo a la tasa general de inflación medida mediante el IPC.
- Costo de universidades públicas.** El costo anual promedio (matrícula, cuotas, alojamiento y comida) en una universidad pública para cuatro años se elevó de \$2 490 en 1980 a \$11 354 en 2004. Calcule el porcentaje de aumento en el costo de 1980 a 2004 y compárelo a la tasa general de inflación medida mediante el IPC.
- Precios de casas en el sur.** El precio típico (medio) de una casa nueva unifamiliar en la parte sur de Estados Unidos se elevó de \$75 300 en 1990 a \$155 500 en 2004. Calcule el porcentaje de aumento en el costo de una casa de 1990 a 2004 y compárelo con la tasa general de inflación medida con el índice de precios al consumidor.
- Precios de casas en el oeste.** El precio típico (medio) de una casa nueva unifamiliar en la parte oeste de Estados Unidos se elevó de \$129 600 en 1990 a \$241 300 en 2004. Calcule el porcentaje de aumento en el costo de una casa de 1990 a 2004 y compárelo con la tasa general de inflación medida con el índice de precios al consumidor.

Índice de precios de vivienda. Los agentes inmobiliarios utilizan un índice para comparar los precios de casas en las ciudades principales en todo el país. Los números índices para varias ciudades se dan en la tabla siguiente. Si conoce el precio de una casa particular en su ciudad, puede utilizar el índice para determinar el precio de una casa comparable en otra ciudad:

$$\frac{\text{precio}}{\text{(otra ciudad)}} = \frac{\text{precio}}{\text{(su ciudad)}} \times \frac{\text{índice en otra ciudad}}{\text{índice en su ciudad}}$$

Utilice el índice de precios de viviendas en los ejercicios 17 al 20.

Ciudad	Índice	Ciudad	Índice
Denver	100	Boston	358
Miami	194	Las Vegas	101
Phoenix	86	Dallas	81
Atlanta	90	Cheyenne	60
Baltimore	150	San Francisco	382

17. **Precios de viviendas.** Para una casa valuada en \$300 000 en Denver, determine el precio de una casa equivalente en Miami y en Cheyenne.
18. **Precios de viviendas.** Para una casa valuada en \$500 000 en Boston, determine el precio de una casa equivalente en Baltimore y en Phoenix.
19. **Precios de viviendas.** Para una casa valuada en \$250 000 en Cheyenne, determine el precio de una casa equivalente en San Francisco y en Boston.
20. **Precios de viviendas.** Para una casa valuada en \$1 000 000 en Boston, determine el precio de una casa equivalente en San Francisco y en Cheyenne.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 2 en www.aw.com/bbt.

21. **Índice de precios al consumidor.** Vaya a la página principal (home page) del índice de precios al consumidor y determine los últimos valores publicados para el IPC. Resuma éstos y cualquier tendencia en el IPC.
22. **Índice de precios al productor.** Vaya a la página principal (home page) del índice de precios al productor (IPP). Lea las noticias generales y más recientes publicaciones. Escriba un breve resumen que describa el propósito del IPP y en que se diferencia del IPC. También resuma cualquier tendencia importante reciente en el IPP.
23. **Índice de confianza del consumidor.** Utilice un motor de búsqueda para determinar noticias recientes acerca del índice de confianza del consumidor. Después de estudiar la noticia, escriba un resumen breve de lo que este índice trata de medir y describa cualquier tendencia reciente del índice.
24. **Índice de desarrollo humano.** El Programa de Desarrollo de las Naciones Unidas por lo regular publica su reporte de desarrollo humano. Una conclusión muy observada de cerca de este reporte es el índice de desarrollo humano (IDH), que mide los logros generales en un país en tres dimensiones básicas de desarrollo humano: esperanza de vida, logros en educación e ingreso ajustado. Encuentre la copia más reciente de este reporte e investigue exactamente cómo se define y es calculado el IDH.
25. **Índice de tiendas de conveniencia.** Vaya a un supermercado local y encuentre los precios de unos cuantos productos básicos, pan, leche, jugo y café. Calcule el costo total de esos artículos. Luego vaya a algunas pequeñas tiendas de conveniencia y encuentre los precios de los mismos artículos. Use el total del supermercado como el valor de referencia, calcule los números índice para las tiendas de conveniencia.
26. **Precios de gasolina.** Recolecte datos, haga una gráfica convincente y escriba un argumento persuasivo, ya sea para defender o refutar el enunciado que los precios de la gasolina *no* han aumentado en los últimos 30 años relativos al costo total de la vida.

EN LAS NOTICIAS

27. **Índice de precios al consumidor.** Determine un artículo reciente que incluya una referencia al IPC. Describa brevemente cómo el IPC es importante en la historia.
28. **Números índices.** Determine un artículo reciente que incluya un número índice distinto al IPC. Describa el número índice y su significado, y analice cómo es importante el índice en la historia.

Ejercicios de repaso del capítulo

1. Titanic. De las 2223 personas que abordaron el *Titanic*, 31.76% sobrevivió cuando el barco se hundió el lunes 15 de abril de 1912.

- Determine el número verdadero de personas que sobrevivió cuando se hundió el *Titanic*.
- ¿El número de sobrevivientes es un valor de un conjunto de datos discretos o de un conjunto de datos continuos? Explique.
- De las personas que abordaron el *Titanic*, 531 eran mujeres o niños. ¿Qué porcentaje de pasajeros del *Titanic* fueron mujeres o niños?
- Había 45 niñas a bordo del *Titanic* y había 42% más niños que niñas. ¿Cuántos niños había a bordo?
- Si compilamos las edades de todos los pasajeros a bordo del *Titanic*, ¿cuál es el nivel de medida (nominal, ordinal, intervalo, razón) de esas edades?
- Si listamos a los pasajeros de acuerdo con las categorías de hombres, mujeres, niños y niñas, ¿cuál es el nivel de medida (nominal, ordinal, intervalo, razón) de este conjunto de datos?

2. Encuesta de AOL. En una encuesta de America OnLine a 3309 personas, 66% dijo que algunas veces comían en KFC, 26% dijo que ellos nunca comían en KFC y 8% dijo que con frecuencia comían en KFC.

- ¿Cuál es el número de encuestados que dijeron que nunca comían en KFC?
- Si 1224 de los encuestados dijo que KFC ni ganaría ni perdería después de la eliminación de las grasas trans, ¿cuál es porcentaje de encuestados que hicieron esta afirmación?
- Dado que las respuestas posibles son *siempre*, *algunas veces* y *nunca*, ¿el nivel de medida de esas respuestas es nominal, ordinal, intervalo o de razón?
- Dado que la encuesta fue llevada a cabo preguntando a los usuarios de America OnLine para responder a una pregunta que fue puesta en el sitio web, ¿qué concluye acerca de los resultados de la encuesta? ¿Es probable que los resultados reflejen la opinión de la población general?

3. Gasto en cuidado de la salud. El total de gasto en cuidado de la salud en Estados Unidos se elevó de \$80000 millones en 1973 a \$1.8 billones en 2004. El índice de precios al consumidor fue 44.4 en 1973 y fue 188.9 en 2004 (con 1982-1984 = 100). Compare el cambio en el gasto en salud gastado de 1973 a 2004 con la tasa general de inflación medida por medio del IPC.

4. Salario mínimo. La tabla siguiente lista los salarios mínimos federales por hora en los pasados 60 años, tanto en dólares reales en el instante como en dólares de 1996. Salvo por 2006, las entradas de la tabla corresponden a años en que el salario mínimo cambió (con base en la información del Departamento del Trabajo).

Año	Dólares reales	Dólares de 1996
1938	0.25	2.78
1939	0.30	3.39
1945	0.40	3.49
1950	0.75	4.88
1956	1.00	5.77
1961	1.25	6.41
1967	1.40	6.58
1968	1.60	7.21
1974	2.00	6.37
1976	2.30	6.34
1978	2.65	6.38
1979	2.90	6.27
1981	3.35	5.78
1990	3.50	4.56
1991	4.25	4.90
1996	4.75	4.75
1997	5.15	5.03
2006	5.15	4.01

- De acuerdo con la tabla, ¿cuál es el valor de \$0.25 en dólares de 1938 en dólares de 1996?
- De acuerdo con la tabla, ¿cuál es el valor de \$1.00 en dólares de 1956 en dólares de 1996?
- ¿Por qué el salario mínimo para 2006 en dólares reales es mayor que el salario mínimo para 2006 en dólares de 1996?

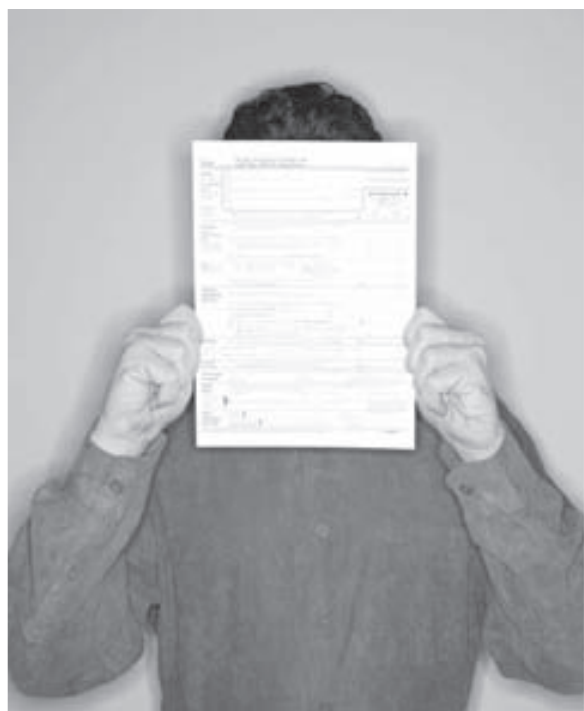
Cuestionario del capítulo

1. El color de ojos de sujetos seleccionados aleatoriamente se registraron como parte de un estudio de salud nacional. ¿El nivel de medida de esos colores de ojos es nominal, ordinal, de intervalo o de razón?
2. Las circunferencias de las cabezas de sujetos elegidos aleatoriamente se registraron como parte de un estudio de salud nacional. ¿Esos valores son continuos o discretos?
3. ¿El nivel de medida de las circunferencias de las cabezas, descritas en el ejercicio 2 es nominal, ordinal, de intervalo o de razón?
4. Un investigador mide la circunferencia de la cabeza de un sujeto y registra un valor de 45.4 centímetros, pero la circunferencia real de la cabeza del sujeto es 55.4. ¿Cuál es el error absoluto?
5. Un investigador mide la circunferencia de la cabeza de un sujeto y registra un valor de 45.4 centímetros, pero la circunferencia real de la cabeza del sujeto es 55.4. ¿Cuál es el error relativo?
6. En una encuesta de Gallup a 1038 adultos, 540 dijeron que fumar colillas de cigarro es muy dañino. ¿Cuál es el porcentaje de adultos que dijeron que fumar colillas de cigarro es muy dañino?
7. En una encuesta de Gallup a 1038 adultos, 5% dijo que fumar colillas de cigarro no es tan dañino. ¿Cuál es el porcentaje de adultos que dijeron que fumar colillas de cigarro no es tan dañino?
8. Dos estudiantes miden la estatura de un instructor que, en realidad, mide 178.44 centímetros de altura. El primer estudiante obtiene una medida de 178 centímetros y el segundo una medida de 179.18 centímetros. ¿Cuál medida es más exacta? ¿Cuál medida es más precisa?
9. La compañía Telektronics ha estado en actividad durante 5 años y la tabla lista las utilidades netas en cada uno de esos años. Con el primer año como referencia, determine el número índice para la utilidad neta en el segundo año.

Año	Utilidad neta
1	\$12,335
2	\$15,257
3	\$23,444
4	\$31,898
5	\$47,296

10. Con respecto a la tabla en el ejercicio 9. Si la utilidad neta en el sexto año está proyectada que será 12% más que en el quinto año, ¿cuál es la utilidad neta proyectada para el sexto año?

HABLEMOS DE POLÍTICA



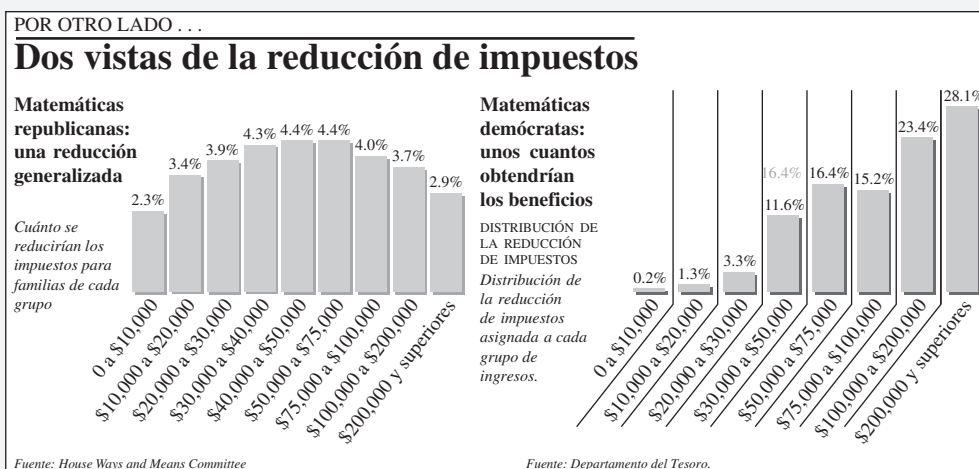
¿Quién se beneficia con una reducción de impuestos?

Los políticos tienen una asombrosa capacidad para proyectar números en la forma que mejor apoyen sus convicciones. Considere los dos diagramas que se muestran en la figura 2.4. Ambos pretenden mostrar los efectos de la *misma* propuesta de reducción de impuestos, sin embargo, parecen respaldar conclusiones radicalmente diferentes. (Esta reducción de impuestos en particular no se convirtió en ley).

El diagrama de la figura 2.4a refleja los números proporcionados por los republicanos, quienes apoyaron la reducción de impuestos. Sugiere que la reducción de impuestos beneficiaría a familias de ingresos similares en términos de porcentaje; con beneficios un poco mejores para familias de medianos ingresos que para familias pobres o ricas. La figura 2.4b refleja los números proporcionados por los demócratas, quienes se opusieron a la reducción de impuestos. Este diagrama sugiere que los beneficios se irían desproporcionadamente hacia los ricos. Si suponemos que ninguna de las partes miente, ¿cómo pueden hacer afirmaciones tan diferentes? La respuesta reside en cómo cada parte elige calcular “los beneficios” de la reducción de impuestos.

Los republicanos calcularon el *promedio* de la reducción de impuestos que sería recibida por las familias de cada grupo. Por ejemplo: la última columna de la figura 2.4a muestra que las familias con ingresos superiores a los \$200 000 podrían tener un promedio de 2.9% en la reducción de impuestos. Por supuesto, una reducción de 2.9% significa mucho más dinero para alguien que paga una tasa alta de impuestos que para alguien que paga menos. Por ejemplo: alguien que paga \$100 000 en impuestos, ahorraría \$2 900 con 2.9% de reducción de impuestos, mientras otra persona que sólo paga \$1 000 en impuestos, ahorraría \$29.

Los demócratas calcularon el porcentaje del *total* de beneficio para cada grupo. Por ejemplo, la última columna de la figura 2.4b muestra que las familias con ingresos superiores a los \$200 000 recibiría 28.1% del total de beneficios de la reducción de impuestos. Pero esto



(a)

(b)

Figura 2.4 Fuente: Adaptada del *New York Times*.

no significa que estas familias conseguirían 28.1% de reducción de impuestos. Para ver por qué, considere los efectos de una reducción de impuestos del mismo porcentaje para cada uno. Puesto que las familias con ingresos superiores a los \$200 000 pagan más de un cuarto del total de impuestos recaudados por el gobierno de Estados Unidos, estas familias recibirían un cuarto del total de los beneficios para *cualquier* reducción generalizada. Por ejemplo, si el gobierno recolectase un billón de dólares de impuestos, 10% significa un ahorro de \$100 mil millones para los contribuyentes, como un todo. En ese caso, ya que un cuarto del billón de dólares fue pagado por familias con ingresos superiores a \$200 000, un cuarto de los \$100 mil millones ahorrados se canalizarían hacia estas mismas familias. De esta manera, estas familias obtendrían 25% de los *beneficios* de la reducción de impuestos aunque su reducción de impuestos real fuese de 10%.

¿Qué partido fue más justo? Ninguno realmente. Los republicanos optaron por no tomar en cuenta el hecho de que la mayoría del total de los impuestos se va hacia los ricos, mientras que los demócratas eligieron despreciar el hecho que los ricos pagan la mayor parte de los impuestos. Lamentablemente este modo de “elegir la verdad” es muy común cuando se trata de números, especialmente aquellos que se emplean en la política.

Políticos y miembros del gobierno con frecuencia abusan de los números y de la lógica en las formas más elementales. Ellos sólo acomodan las cifras para que convengan a sus propuestas, usan medidas vagas de desempeño económico y dan ejemplos tremendos de abuso de gráficas, todo en nombre de disfrazar verdades difíciles de aceptar.

—A.K. Dewdney,
200% of Nothing

PREGUNTAS PARA DISCUSIÓN

1. Los porcentajes de la figura 2.4a muestran el ahorro *relativo* de impuestos para cada grupo de ingresos. ¿Cómo compara estos ahorros relativos con los ahorros *absolutos* de cada grupo? (Sugerencia: al estimar el ahorro absoluto en impuestos, recuerde que la cantidad de impuestos que paga cada familia es algún porcentaje de sus ingresos y que, por lo general, las familias con bajos ingresos pagan, en impuestos, un porcentaje menor de sus ingresos).
2. Una razón de menor importancia en las diferencias en ambas gráficas es que los dos partidos definen *ingresos* de diferente manera. Por ejemplo: los demócratas decidieron asignar las ganancias de las empresas a los ingresos individuales de los accionistas, esto significa que una persona, tenedora de acciones, fue catalogada con mayores ingresos por los demócratas que por los republicanos. ¿Cómo la diferente definición de ingresos afecta a las dos gráficas?
3. ¿Cree usted que alguna de las gráficas en la figura 2.4 describe con precisión “la justicia” global de la iniciativa de reducción de impuestos? Si es así, ¿cuál gráfica y por qué? Y si dice que no, ¿cómo cree usted que los números podrían describirse de manera más justa?
4. ¿Alguna reducción o incremento de impuestos ha sido propuesta por el Congreso de Estados Unidos o por el presidente este año?

LECTURAS SUGERIDAS

“Two Views of a Tax Cut”, *New York Times*, 7 de abril de 1995.

Andrews Edward, “Your Taxes: Cracking the Tax Code”, *New York Times*, 11 de febrero de 2006.

Andrews Edward, “Tax Cuts Offer Most for Very Rich, Study Says”, *New York Times*, 8 de enero 2007.

HABLEMOS DE ECONOMÍA



¿Está mejorando nuestro nivel de vida?

A la mayoría de nosotros nos gustaría que nuestro nivel de vida estuviese mejorando cada año, y en principio es fácil decir si ése es el caso: si su ingreso se eleva más rápido que la tasa de inflación, entonces su nivel de vida se elevará, si su ingreso no se mantiene arriba de la inflación, entonces su nivel de vida caerá.

Sin embargo, en realidad, la situación es mucho más compleja. Por ejemplo, incluso si sus ingresos se elevan sustancialmente, su nivel de vida podría caer, si encara nuevos gastos, tal como una crisis de salud que su seguro no cubrirá o costos propios o un hijo que va a la universidad, o un pago de hipoteca de tasa variable que se ajusta a un costo mensual más alto.

Complicaciones similares afectan la pregunta del nivel de vida nacional. Los economistas coinciden, casi por unanimidad, que el nivel de vida del estadounidense promedio se ha elevado de manera sustancial a partir de la Segunda Guerra Mundial hasta principios de la década de los setenta, y luego nuevamente durante el final de la década de los noventa. Pero, a menos que usted esté en el pequeño porcentaje con salarios altos (cuyo ingreso se ha elevado de manera continua; vea “Hablemos de economía; ¿Los ricos se vuelven más ricos?” al final del capítulo 4), existe un debate considerable sobre los niveles de vida durante las décadas de los setenta y ochenta y durante los primeros años del nuevo milenio.

La pregunta clave en el debate es si el *ingreso real* —ingreso ajustado por la inflación— está subiendo o bajando. Esta pregunta, a su vez, depende de cómo medimos la inflación. Como se analizó en la sección 2.4, la inflación por lo común se mide con el índice de precios al consumidor (IPC). La figura 2.5 muestra cómo los salarios reales de

la mayoría de los estadounidenses han cambiado desde la década de los ochenta, si utilizamos el IPC como la medida de la inflación. Por ejemplo, un cambio en los salarios reales de 0% significa que los salarios reales y el IPC cambiaron exactamente en la misma cantidad en un año particular. El cambio en los salarios reales es positivo cuando los salarios reales aumentan más que el IPC y es negativo cuando los salarios reales aumentaron menos que el IPC.

Observe que, de acuerdo con esta medida, los salarios reales han disminuido (cambios negativos) en la mayoría de los años desde finales de la década de los setenta. En realidad, si calcula el cambio total durante el periodo mostrado (y regresamos unos cuantos años más), esta figura indica que los salarios reales en 2006 fueron *menores* que los que se tenían tres décadas antes. En las elecciones de 2006 estos datos ayudaron a los demócratas que competían por el Congreso a asegurar que los salarios —y por tanto, el nivel de vida— para la mayoría de los estadounidenses se habían estancado durante más de 30 años.

¿Pero es la imagen de que nuestro nivel de vida en realidad es desolador? ¿O en realidad es aún más desolador? La respuesta depende de a quién le pregunte, ya que diferentes economistas tienen opiniones diferentes de cómo el IPC exagera, minimiza o mide de manera precisa la inflación.

Iniciemos con aquellos que argumentan que el IPC *calcula en exceso* la inflación. Si ese es el caso, entonces la inflación real es *menor* a lo que indica el IPC, lo cual significa que la línea cero está demasiado arriba en la figura 2.5. Por ejemplo, si el IPC exagera la tasa de inflación en un punto porcentual cada año, entonces la línea del cero debe estar en el nivel de -1 en la figura 2.5, en lugar de estar en 0. Ese cambio haría positivos la mayoría de los años mostrados en la gráfica en lugar de negativos, lo cual significaría que en lugar de estar estancado el nivel de vida promedio en realidad se habría elevado durante las pasadas tres décadas.



FIGURA 2.5 El cambio en los salarios, ajustados al IPC, para trabajadores no patrones en el sector privado. Estos trabajadores representan alrededor de 80% de la fuerza de trabajo, sin contar a los empleados del gobierno. Los cambios negativos representan disminución en los salarios reales y cambios positivos representan aumento en los salarios reales. Adaptado del *New York Times*, 3 de enero de 2007. Fuente: Departamento del Trabajo de Estados Unidos, Universidad de Michigan, Oficina Nacional de Investigación Económica.

La clave de este argumento, documentado más claramente en el Reporte Boskin de 1996 (vea la lista de lecturas sugeridas), es que el cálculo estándar del IPC tiene errores sistemáticos que aumentan el costo de la vida. El reporte señala dos errores sistemáticos principales que parecen tener este efecto. Primero, de un mes al siguiente, el IPC tiene como base los cambios en los precios de artículos particulares en tiendas particulares. Sin embargo, en realidad, si el precio de un artículo aumenta en una tienda, con frecuencia los consumidores compran más barato en otra tienda. Si el precio de un artículo sube demasiado, los consumidores pueden sustituirlo por uno similar pero de menor precio, lo cual puede significar algo tan simple como un cambio de marcas. Por tanto, los consumidores no encuentran que sus costos reales se hayan elevado tanto como indica el IPC. Segundo, el IPC sigue la pista de artículos comprados por consumidores “típicos” en cualquier instante dado. Estos artículos cambian con el tiempo, en especial cuando provienen de tecnología, pero el IPC no toma en cuenta que tales cambios afectan nuestro nivel de vida. Por ejemplo, hace 20 años nadie era propietario de un grabador de DVD, un iPod, una televisión de alta definición o una computadora con acceso a internet de alta velocidad. La mayoría de los economistas dirían que estos artículos han mejorado nuestro nivel de vida, pero en el IPC sólo cuentan como artículos que actualmente los consumidores “típicos” compran. Con base en estos tipos de errores sistemáticos, el Reporte Boskin concluyó que el IPC sobreestima la tasa real de inflación entre 0.8 y 1.6 puntos porcentuales por año.

Aunque la mayoría de los economistas reconocen los problemas identificados por el Reporte Boskin, algunos argumentan que otros errores sistemáticos funcionan en la dirección opuesta. Por ejemplo, muchos trabajadores de mediana edad ahora tienen gastos significativos asociados con el cuidado de sus padres en la vejez, algo que no fue cierto durante décadas, y en promedio todos pagamos un porcentaje mucho más alto de ingresos en seguro de salud y atención médica de lo que pagábamos en el pasado. Estos tipos de gastos dejan menos ingresos para el tipo de artículos que, por lo general, consideramos que mejoran nuestro nivel de vida. En realidad, algunos economistas argumentan que estos cambios han sido tan significativos que el IPC en realidad *subestima* la verdadera tasa de inflación, en cuyo caso nuestros niveles de vida han caído más de lo que indica la gráfica de los salarios en la figura 2.5.

La lección en este estudio debe ser clara. Las estadísticas que consideran al IPC y los cambios en los salarios reales son muy claras, aunque personas de diferentes creencias políticas o sistemas políticos puedan sacar conclusiones muy diferentes de estos datos. Después de todo, eso es por lo cual la gente puede no estar de acuerdo, de buena fe, acerca de interrogantes como si nuestro nivel de vida está subiendo o bajando. También es por lo que debemos pensar cuidadosamente acerca de las afirmaciones estadísticas de todo tipo, en lugar de sólo aceptar el último eslogan político sin cuestionarlo.

Si tuviese que llenar un manicomio con gente certificada como loca, yo sólo seleccionaría de todas aquellas que aseguran entender la inflación.

—Will Rogers

PREGUNTAS PARA DISCUSIÓN

1. Con base en la gráfica de la figura 2.5, describa en términos generales cómo han cambiado los salarios reales en las pasadas tres décadas. Por ejemplo, describa el periodo general en el que los salarios se han elevado y aquéllos en que han disminuido.
2. En general, ¿considera que la figura 2.5 proporciona una representación exacta de los cambios en el nivel de vida para los estadounidenses promedio? Defienda su opinión.
3. Considere los argumentos en el artículo que sugieren errores sistemáticos que podrían provocar que el IPC sobreestimara o subestimara la inflación. Seleccione uno de estos errores sistemáticos e investiguelo con mayor detalle. Proporcione suficientes ejemplos, saque su propia conclusión acerca del cambio en nuestro nivel de vida.
4. En la figura 2.5 observe el drástico aumento en los salarios reales que se muestra para 2006. Haga una búsqueda en la web para determinar qué sucedió a los salarios desde ese momento. ¿Ha continuado el aumento o sólo fue temporal?

LECTURAS SUGERIDAS

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches y Dale W. Jorgenson, "Toward a More Accurate Measure of the Cost of Living", *Informe final al Comité de Finanzas del Senado, de la Comisión Asesora para el Estudio del Índice de Precios al Consumidor*, Washington, DC, Comité de Finanzas del Senado, 1996.

Greenhouse, Steven, "Falling Fortunes of Wage Earners", *New York Times*, 12 de abril de 2005.

Gross, Jane, "Elder-Care Costs Deplete Savings of a Generation", *New York Times*, 30 de diciembre de 2006.

Leonhar, David, "It's the Year to Keep an Eye on Paychecks", *New York Times*, 3 de enero de 2007.



Lo más grandioso de una pintura es cuando nos fuerza a observar lo que nunca esperábamos ver.

—John Tukey

Exhibición visual de datos

SI USTED VE UN PERIÓDICO, UN REPORTE ANUAL DE una empresa o un estudio gubernamental, casi seguramente verá tablas y gráficas de datos estadísticos. Algunas de estas tablas y gráficas son muy sencillas, otras pueden ser muy complejas. Unas facilitan entender los datos; otras pueden ser confusas o incluso engañosas. En este capítulo estudiaremos las formas principales en que los datos estadísticos se muestran en tablas y gráficas. Puesto que la capacidad para transmitir conceptos por medio de gráficas es tan valiosa en la actual sociedad de la información, las habilidades desarrolladas en este capítulo son muy importantes para el éxito en casi cualquier profesión.

OBJETIVOS DE APRENDIZAJE

3.1 Tablas de frecuencia

Ser capaz de crear e interpretar tablas de frecuencia.

3.2 Distribución gráfica de datos

Ser capaz de crear e interpretar gráficas de barras, diagramas de puntos, gráficas circulares, histogramas, diagramas de tallos y hojas, gráficas de líneas y diagramas de series de tiempo.

3.3 Gráficas en los medios

Entender cómo interpretar los diversos tipos de gráficas más complejas que se encuentran regularmente en noticias de los medios.

3.4 Algunas precauciones con respecto a gráficas

Evaluar de una manera crítica las gráficas e identificar formas comunes en las que las gráficas pueden ser engañosas.

3.1 Tablas de frecuencia

Tabla 3.1 Tabla de frecuencias para un conjunto de calificaciones de un ensayo

Calificación	Frecuencia
A	4
B	7
C	9
D	3
F	2
Total	25

La profesora Delaney registra la lista siguiente de calificaciones de sus 25 estudiantes en un conjunto de ensayos:

A C C B C D C C F D C C C
B B A B D B A A B F C B

Esta lista tiene todas las calificaciones, pero no es muy fácil de leer. Una forma mucho más sencilla de mostrar los datos es mediante la construcción de una tabla en la que registramos el número de veces, o **frecuencia**, que apareció cada calificación. El resultado, que se muestra en la tabla 3.1, se denomina **tabla de frecuencias**. Las cinco posibles calificaciones (A, B, C, D, F), se denominan **categorías** (o clases) para la tabla.

Definición

Una **tabla básica de frecuencias** tiene dos columnas:

- Una columna lista todas las **categorías** de los datos.
- La otra columna lista la **frecuencia** de cada categoría, que es el número de valores de los datos en la categoría.

EJEMPLO 1 Prueba de sabor

La compañía de bebidas Rocky Mountain quiere conocer la reacción a su producto nuevo, Coral Cola, y prepara una prueba de sabor con 20 personas. A cada individuo se le pide que clasifique el sabor de la cola en una escala de 5 puntos:

(mal sabor) 1 2 3 4 5 (excelente sabor)

Tabla 3.2 Calificaciones de prueba de sabor

Escala de sabor	Frecuencia
1	2
2	3
3	9
4	4
5	2
Total	20

Las 20 calificaciones son las siguientes:

1 3 3 2 3 3 4 3 2 4
2 3 5 3 4 5 3 4 3 1

Construya una tabla de frecuencias para estos datos.

Solución La variable de interés es *sabor* y esta variable puede tomar cinco valores: las categorías de sabor de 1 a 5. (Observe que los datos son cualitativos y están en el nivel de medida ordinal). Construimos la tabla con estas cinco categorías en la columna izquierda y sus frecuencias en la columna derecha, como se muestra en la tabla 3.2.

Clasificación de datos en clases

Considere los datos en la tabla 3.3, que muestra el promedio anual de energía que consume una persona, en millones de BTU, en cada uno de los 50 estados de la Unión Americana. Los 50 números en este conjunto de datos (sin contar el promedio de Estados Unidos) varía de 212 millones de BTU por persona (Rhode Island) hasta 1 175 millones de BTU por persona (Alaska), y la mayoría de los números sólo aparecen una vez. ¿Cómo podemos, de manera eficiente, construir una tabla de frecuencias de estos datos?

La respuesta es crear categorías que cubran algún rango de valores de los datos. Por ejemplo, podríamos crear una categoría para todos los valores entre 200 y 299 millones de

Tabla 3.3 Energía anual promedio usada por persona, por estado, en millones de BTU

Estado	Millones de BTU por persona	Estado	Millones de BTU por persona	Estado	Millones de BTU por persona
Alabama	447	Louisiana	822	Ohio	349
Alaska	1175	Maine	366	Oklahoma	425
Arizona	246	Maryland	281	Oregon	295
Arkansas	415	Massachusetts	248	Pennsylvania	321
California	229	Michigan	313	Rhode Island	212
Colorado	297	Minnesota	355	South Carolina	389
Connecticut	255	Mississippi	411	South Dakota	345
Delaware	383	Missouri	322	Tennessee	388
Florida	252	Montana	410	Texas	560
Georgia	343	Nebraska	372	Utah	296
Hawaii	248	Nevada	292	Vermont	252
Idaho	341	New Hampshire	254	Virginia	329
Illinois	310	New Jersey	298	Washington	316
Indiana	470	New Mexico	353	West Virginia	433
Iowa	400	New York	220	Wisconsin	335
Kansas	410	North Carolina	314	Wyoming	919
Kentucky	456	North Dakota	624	Promedio de Estados Unidos	339

Nota: Los datos incluyen todos los usos de energía, incluyendo residencial, comercial, industrial y transporte.

Fuente: Administración de Energía, Estados Unidos (tablas de 2007, con información de 2003).

BTU, luego crear una segunda categoría para los valores entre 300 y 399 millones de BTU, y así sucesivamente. Luego contamos la frecuencia (número de valores) en cada categoría, generando la tabla de frecuencia que se muestra como tabla 3.4. Este proceso se denomina **clasificación en clases** de los datos, ya que cada categoría actúa como una clase separada en la que podemos colocar algunos de los datos.

Definición

Cuando es imposible o poco práctico tener una categoría para cada valor en un conjunto de datos, **clasificamos** (o agrupamos) los datos en categorías (*clases*), cada una cubriendo un rango de posibles valores.

Tabla 3.4 Tabla de frecuencias para la información del uso de energía de la tabla 3.3

Uso anual de energía por persona (millones de BTU)	Frecuencia (número de estados)
200-299	16
300-399	19
400-499	10
500-599	1
600-699	1
700-799	0
800-899	1
900-999	1
1000-1099	0
1100-1199	1
Total	50

A propósito...

Una BTU, o unidad térmica británica, es la energía necesaria para elevar la temperatura de una libra de agua en 1°F. Es equivalente a 252 calorías o 1 055 joules.

Tabla 3.5 Datos para acciones en el promedio industrial Dow Jones, 2006

Compañía	Ingreso (miles de millones)	Rendimiento*	Rango**	Compañía	Ingreso (miles de millones)	Rendimiento*	Rango**
Wal-Mart	351.1	0.1%	1	Johnson & Johnson	53.3	12.4%	36
Exxon Mobil	347.3	39.2%	2	Pfizer	52.4	15.3%	39
General Motors	207.3	64.2%	3	United Technologies	47.8	13.7%	42
General Electric	168.3	26.5%	5	Microsoft	44.3	15.8%	49
Citigroup	146.8	19.4%	8	Caterpillar	41.6	7.9%	55
AIG	113.2	6.0%	10	Intel	35.4	-17.2%	62
J.P. Morgan Chase	100.0	25.5%	11	Walt Disney	34.3	44.5%	64
Verizon	93.2	34.6%	13	Honeywell	31.4	24.2%	69
Hewlett-Packard	91.7	45.3%	14	ALCOA	30.9	3.5%	71
IBM	91.4	19.8%	15	DuPont	29.0	18.7%	74
Home Depot	90.8	0.9%	17	American Express	27.1	19.1%	79
Altria	70.3	19.9%	23	Coca-Cola	24.1	23.1%	94
Procter & Gamble	68.2	13.3%	25	3M	22.9	3.0%	97
AT&T	63.1	52.9%	27	Merck	22.6	42.9%	99
Boeing	61.5	28.4%	28	McDonald's	21.6	34.6%	108

*Rendimiento total para inversionistas, 2006, incluyendo dividendos y cambio en el precio de la acción.
**Rango, con base en el ingreso entre 500 compañías de la revista *Fortune*.
Fuente: Fortune.com

Apropósito...

Las 30 acciones que conforman la lista Dow son elegidas por los editores de *Wall Street Journal*. En ocasiones las acciones de la lista son cambiadas. Por ejemplo, en abril de 2004, se quitaron AT&T, Eastman Kodak e International Paper, y fueron reemplazadas por AIG, Pfizer y Verizon. AT&T pronto regresó al índice cuando otra compañía de la lista (SBC Communications) se fusionó con ella y tomó el nombre AT&T.

EJEMPLO 2 Las acciones Dow

Para las 30 acciones del promedio industrial Dow Jones, la tabla 3.5 muestra el ingreso anual (en miles de millones de dólares), el rendimiento total de un año y el rango en la lista de 500 compañías con mayores ingresos en Estados Unidos de la revista *Fortune*. Analice las ventajas y las desventajas de la elección de clases.

Solución Los datos de ingresos varían desde \$21.6 mil millones (McDonald's) a \$351.1 mil millones (Wal-Mart). Existen muchas posibles maneras de clasificar los datos para este rango; a continuación está una buena y la razón para ello:

- Creamos clases que cubran un rango de \$0 a \$400 mil millones. Eso cubre el rango completo de los datos, con espacio adicional para los valores más pequeños y por arriba para los valores más altos.
- A cada clase le damos un ancho de \$50 mil millones, de modo que podemos cubrir el rango de \$0 a \$400 mil millones con ocho clases. Además, el ancho de \$50 mil millones es un número adecuado que ayuda a construir una tabla fácil de leer.
- Puesto que los datos están dados al décimo (de mil millón) más próximo, también definimos las clases al décimo más cercano, así que no se traslapan. Esto es, las clases van de \$0 a \$49.9 mil millones, de \$50.0 a \$99.9 mil millones, y así sucesivamente.

La tabla 3.6 muestra las frecuencias resultantes.

Tabla 3.6 Tabla de frecuencias para la información de ingreso anual de la tabla 3.5

Ingreso anual (miles de millones de dólares)	Frecuencia (número de compañías)
0-49.9	13
50-99.9	10
100-149.9	3
150-199.9	1
200-249.9	1
250-299.9	0
300-349.9	1
350-399.9	1
Total	30

UN MOMENTO DE REFLEXIÓN

Considere otras tres posibilidades de clasificar los datos en la tabla 3.6: 4 clases que cubran el rango de \$0 a \$400 mil millones, 11 clases que cubran el rango de \$0 a \$375 mil millones y 36 clases que cubran de \$0 a \$360 mil millones. Analice brevemente las ventajas y las desventajas de cada una de estas elecciones.

Frecuencia relativa

Vuelva a considerar las calificaciones del ensayo que se listan en la tabla 3.1. Podríamos necesitar saber no sólo el número de estudiantes que obtuvieron cada calificación, sino también la fracción o porcentaje de estudiantes que obtuvieron esas calificaciones. A estas fracciones (o porcentajes) para cada categoría les llamamos **frecuencias relativas**. Por ejemplo, 4 de los 25 estudiantes obtuvieron calificación de A, así que la frecuencia relativa de A es $4/25$, o 0.16, o bien 16%. La tabla 3.7 repite la información de la tabla 3.1, pero esta vez con una columna adicional para la frecuencia relativa.

Tabla 3.7 Tabla de frecuencias relativas		
Calificación	Frecuencia	Frecuencia relativa
A	4	$4/25 = 0.16$
B	7	$7/25 = 0.28$
C	9	$9/25 = 0.36$
D	3	$3/25 = 0.12$
F	2	$2/25 = 0.08$
Total	25	1

La suma de las frecuencias relativas debe ser igual a 1 (o 100%), ya que cada frecuencia relativa es una fracción de la frecuencia total. (En ocasiones, los redondeos provocan que el total sea un poco diferente de 1).

*“¡Datos! ¡Datos! ¡Datos!”
exclamó de manera impaciente.
“No puedo hacer ladrillos sin
arcilla.”*

—Sherlock Holmes en *La aventura de la finca de Cooper Beeches*
de Sir Arthur Conan Doyle

Definición

La **frecuencia relativa** de cualquier categoría es la proporción o porcentaje de los datos que caen en esa categoría:

$$\text{frecuencia relativa} = \frac{\text{frecuencia de la categoría}}{\text{frecuencia total}}$$

Frecuencia acumulada

Observe una vez más las calificaciones del ensayo en la tabla 3.1. ¿Qué pasa si quiere conocer cuántos estudiantes obtuvieron una calificación de C o mejor? Por supuesto, podríamos sumar las frecuencias de las calificaciones A, B y C para encontrar que 20 estudiantes obtuvieron una C o mejor calificación. Sin embargo, con frecuencia, las tablas hacen esta aritmética por nosotros, mostrando las **frecuencias acumuladas** o el número de datos en una categoría en particular y *todas las categorías que le preceden*. La tabla 3.8 repite la información de la tabla 3.1, pero esta vez con una columna adicional para la frecuencia acumulada. Observe que la frecuencia acumulada para la última categoría siempre debe ser igual al número total de datos, que es la frecuencia total.

NOTA TÉCNICA

La mayoría de las tablas de frecuencias inician con la categoría más baja, pero las tablas de calificaciones por lo común inician con la categoría más alta (A). El orden de las categorías no afecta las frecuencias ni las frecuencias relativas, pero *sí* afecta las frecuencias acumulativas. Por ejemplo, si la tabla de la derecha tiene las categorías en orden inverso, la frecuencia acumulativa para C sería el número de calificaciones de C o menores (en lugar de C o mayores).

Tabla 3.8 Tabla de frecuencias acumuladas

Calificación	Frecuencia	Frecuencia acumulada
A	4	4
B	7	7 + 4 = 11
C	9	9 + 7 + 4 = 20
D	3	3 + 9 + 7 + 4 = 23
F	2	2 + 3 + 9 + 7 + 4 = 25
Total	25	25

Definición

La **frecuencia acumulada** de cualquier categoría es el número de datos en esa categoría y en todas las categorías precedentes.

Tenga en cuenta que las frecuencias acumuladas sólo tienen sentido para datos categóricos que tienen un orden claro. Es decir, podemos utilizar frecuencias acumulativas para datos en los niveles de medida ordinal, de intervalo y de razón, pero no podemos hacerlo en datos en el nivel de medida nominal.

EJEMPLO 3 Más sobre la prueba de sabor

Mediante los datos de la prueba de sabor del ejemplo 1, construya una tabla de frecuencias con columnas para las frecuencias relativa y acumulada. ¿Qué porcentaje de los encuestados dieron a la cola la calificación más alta? ¿Qué porcentaje dio a la cola una de las tres calificaciones más bajas?

Solución Determinamos las frecuencias relativas dividiendo la frecuencia de cada categoría entre la frecuencia total de 20. Encontramos las frecuencias acumuladas sumando las frecuencias en cada categoría a la suma de las frecuencias de todas las categorías que le preceden. La tabla 3.9 muestra los resultados. La columna de frecuencias relativas muestra que 0.10, o 10%, de los encuestados dio a la cola la calificación más alta. La columna de frecuencia acumulada muestra que 14 de 20 personas, o 70%, dio a la cola una calificación de 3 o menor.

NOTA TÉCNICA

Una frecuencia acumulada, dividida entre la frecuencia total, se denomina *frecuencia relativa acumulada*. Por ejemplo, en la tabla 3.9 la frecuencia relativa acumulada de 3 o menos es $14/20 = 0.70$, lo que significa que 70% de los datos están en las categorías 3, 2 o 1.

Tabla 3.9 Frecuencias relativa y acumulada

Escala de sabor	Frecuencia	Frecuencia relativa	Frecuencia acumulada
1	2	$2/20 = 0.10$	2
2	3	$3/20 = 0.15$	3 + 2 = 5
3	9	$9/20 = 0.45$	9 + 3 + 2 = 14
4	4	$4/20 = 0.20$	4 + 9 + 3 + 2 = 18
5	2	$2/20 = 0.10$	2 + 4 + 9 + 3 + 2 = 20
Total	20	1	20

EJEMPLO 4 Datos de energía

Para la tabla de frecuencias de la energía que se muestra como tabla 3.4, agregue columnas para las frecuencias relativa y acumulada. Analice cualquier tendencia que parezca particularmente reveladora o sorprendente.

Solución Encontramos las frecuencias relativas dividiendo la frecuencia en cada categoría entre la frecuencia total, que en este caso es 50. Las frecuencias acumuladas las encontramos sumando la frecuencia en cada categoría a la suma de las frecuencias en todas las categorías precedentes. La tabla 3.10 muestra los resultados.

La tabla revela muchos hechos interesantes acerca del uso de energía anual por persona en los diferentes estados. Por ejemplo, la columna de frecuencia relativa nos dice que 38% (o 0.38) de los estados están en la categoría de 300-399 millones de BTU por persona por año. La columna de frecuencia acumulada muestra que 35 de 50 estados caen en las primeras dos categorías, o uso de energía menor a 400 millones de BTU por persona por año. Los tres estados con el uso más alto de energía anual por persona utilizan más del doble de energía por persona que cualquiera de los 35 estados con consumo más bajo. (Con la terminología que estudiaremos en el capítulo 4, decimos que estos tres estados tienen valores atípicos, ya que sus valores difieren mucho de los de los otros estados).

Uso anual de energía por persona (millones de BTU)	Frecuencia	Frecuencia relativa	Frecuencia acumulada
200-299	16	0.32	16
300-399	19	0.38	35
400-499	10	0.20	45
500-599	1	0.02	46
600-699	1	0.02	47
700-799	0	0.00	47
800-899	1	0.02	48
900-999	1	0.02	49
1000-1099	0	0.00	49
1100-1199	1	0.02	50
Total	50	1	50

UN MOMENTO DE REFLEXIÓN

Tenga cuidado al interpretar la tabla 3.10 del ejemplo 4. Por ejemplo, la frecuencia relativa 0.32 para 200-290 millones de BTU significa que 32% de los 50 estados tienen un uso de energía per cápita en este rango. ¿Esto también significa que 32% de todos los estadounidenses utilizan entre 200 y 299 millones de BTU cada año? ¿Por qué sí o por qué no? (Sugerencia: vuelva a ver la tabla 3.3 y considere las poblaciones de estos estados).

A propósito...

Las partes de la energía total que van a uso residencial, comercial, industrial y transporte varían ampliamente de un estado a otro. Los usos industrial y de transporte son muy altos en los tres estados con el mayor uso de energía total (Alaska, Wyoming y Louisiana).



Sección 3.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Tabla de frecuencias.** ¿Qué es una tabla de frecuencias? Explique qué queremos decir con categorías (o clases) y frecuencias.
- Frecuencia relativa.** Una tabla de frecuencias tiene cinco clases con frecuencias de 2, 9, 14, 12 y 3. ¿Cuáles son las frecuencias relativas de las cinco clases?
- Frecuencia acumulada.** Una tabla de frecuencia tiene cinco clases con frecuencias de 2, 9, 14, 12 y 3. ¿Cuáles son las frecuencias acumuladas de las cinco clases?
- Tabla de frecuencias.** La primera clase en una tabla de frecuencias muestra una frecuencia de 12, correspondiente al rango de valores de 15.0 a 15.9. Usando únicamente esta información acerca de la tabla de frecuencias, ¿es posible identificar los 12 valores muestrales originales que están resumidos en esta clase? Explique.

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Tabla de frecuencias.** Una amiga le dice que su tabla de frecuencias tiene dos columnas con etiquetas *Estado e Ingreso medio*.
- 6. Frecuencia relativa.** La frecuencia relativa de la categoría A en una tabla es 1.4.
- 7. Frecuencia acumulada.** La frecuencia acumulada de una categoría en una tabla es 1.4.
- 8. Clases.** Para un conjunto de datos, conforme el ancho de las clases disminuye, el número de clases aumenta.

Conceptos y aplicaciones

- 9. Práctica con tabla de frecuencias.** La profesora Díaz registra las calificaciones finales siguientes en uno de sus cursos:

A A A A B B B B B B B C
C C C C C C C D D D F F

- Construya una tabla de frecuencias para estas calificaciones. Incluya columnas para la frecuencia relativa y para la frecuencia acumulada. Explique brevemente el significado de cada columna.
- 10. Práctica con tabla de frecuencias.** Una guía de la ciudad de Nueva York lista restaurantes de 5 estrellas (la más alta calificación), 10 de cuatro estrellas, 20 de tres estrellas, 15 de dos estrellas y 5 de una estrella. Construya una tabla de frecuencias para estas calificaciones. Incluya columnas para las frecuencias relativa y acumulada. Explique brevemente el significado de cada columna.
 - 11. Pesos de Coca.** Construya una tabla de frecuencias para los pesos (en libras) dados a continuación de 36 latas de Coca regular. Inicie la primera clase en 0.7900 libras y utilice un ancho de la clase de 0.0050 libras. Analice sus hallazgos.

0.8192 08150 0.8163 0.8211 0.8181 0.8247
0.8062 08128 0.8172 0.8110 0.8251 0.8264
0.7901 08244 0.8073 0.8079 0.8044 0.8170
0.8161 08194 0.8189 0.8194 0.8176 0.8284
0.8165 08143 0.8229 0.8150 0.8152 0.8244
0.8207 08152 0.8126 0.8295 0.8161 0.8192

- 12. Pesos de Coca de dieta.** Construya una tabla de frecuencias para los pesos (en libras) dados a continuación de 36 latas de Coca de dieta. Inicie la primera clase en 0.7750 libras y utilice un ancho de la clase de 0.0050 libras. Analice sus hallazgos.

0.7773 07758 0.7896 0.7868 0.7844 0.7861
0.7806 07830 0.7852 0.7879 0.7881 0.7826
0.7923 07852 0.7872 0.7813 0.7885 0.7760
0.7822 07874 0.7822 0.7839 0.7802 0.7892
0.7874 07907 0.7771 0.7870 0.7833 0.7822
0.7837 07910 0.7879 0.7923 0.7859 0.7811

- 13. Actores ganadores de un Oscar.** Los datos siguientes muestran las edades de actores hombres ganadores del premio de la Academia en el momento en que ganaron su premio. Construya una tabla de frecuencias para los datos, utilizando clases de 20-29, 30-39 y así sucesivamente. Analice sus hallazgos.

32 37 36 32 51 53 33 61 35 45
55 39 76 37 42 40 32 60 38 56
48 48 40 43 62 43 42 44 41 56
39 46 31 47

- 14. Temperaturas corporales.** Los datos siguientes muestran las temperaturas (°F) de sujetos elegidos aleatoriamente. Construya una tabla de frecuencias con siete clases: 96.9-97.2, 97.3-97.6, 97.7-98.0, etcétera.

98.6 98.6 98.0 98.0 99.0 98.4 98.4 98.4
98.4 98.6 98.6 98.8 98.6 97.0 97.0 98.8
97.6 97.7 98.8 98.0 98.0 98.3 98.5 97.3
98.7 97.4 98.9 98.6 99.5 97.5 97.3 97.6
98.2 99.6 98.7 99.4 98.2 98.0 98.6 98.6
97.2 98.4 98.6 98.2 98.0 97.8 98.0 98.4
98.6 98.6

- 15. Información perdida.** La tabla siguiente muestra las calificaciones de un examen final en una clase de inglés. La tabla está incompleta. Utilice la información dada para llenar las entradas perdidas en la tabla.

Categoría	Frecuencia	Frecuencia relativa
A	?	?
B	?	18%
C	?	24%
D	11	?
F	6	?
Total	50	?

- 16. Información perdida.** La tabla siguiente muestra las calificaciones de dos presentaciones en una clase de teatro. La tabla está incompleta. Utilice la información dada para llenar las entradas perdidas en la tabla.

Categoría	Frecuencia	Frecuencia acumulada
A	?	1
B	6	?
C	7	?
D	?	23
F	?	25
Total	?	?

- 17. Un dado cargado.** Uno de los autores hizo un agujero en un dado, lo llenó con un plomo pesado y luego procedió a lanzarlo. Los resultados se dan en la siguiente tabla de frecuencias.

- De acuerdo con los datos, ¿cuántas veces se tiró el dado?
- ¿Cuántas veces el resultado fue mayor que 2?
- ¿Qué porcentaje de los resultados fueron 6?
- Liste las frecuencias relativas que correspondan a las frecuencias dadas.
- Liste las frecuencias acumuladas que correspondan a las frecuencias dadas.

Resultado	Frecuencia
1	27
2	31
3	42
4	40
5	28
6	32

18. Interpretación de los datos familiares. Considere la tabla de frecuencias para el número de hijos en familias estadounidenses.

- De acuerdo con los datos, ¿cuántas familias hay en Estados Unidos?
- ¿Cuántas familias tienen dos o menos hijos?
- ¿Qué porcentaje de familias estadounidenses no tienen hijos?
- ¿Qué porcentaje de familias estadounidenses tienen tres o más hijos?

Número de hijos	Número de familias (millones)
0	35.54
1	14.32
2	13.28
3	5.13
4 o más	1.97

19. Teclados de computadora. La configuración tradicional del teclado se denomina teclado *Qwerty* por la posición de las letras QWERTY en la fila superior de las letras. Desarrollada en 1872, la configuración *Qwerty* supuestamente forzó a las personas a teclear más despacio de modo que las primeras máquinas de escribir no se atorarán. El teclado *Dvorak*, desarrollado en 1936, supuestamente proporciona un acomodo más eficiente, con la colocación de las teclas más utilizadas en la fila de en medio (o fila “principal”), donde son más accesibles.

Un artículo de la revista *Discover* sugirió que usted podía medir lo fácil de teclear usando el sistema de calificación siguiente: cuente cada letra en la fila principal como 0, cuente cada letra en la fila superior como 1 y cuente cada letra en la fila inferior como 2. Por ejemplo, la palabra *statistics* resultaría en una calificación de 7 en el teclado

Qwerty y de 1 en el teclado *Dvorak*, como se muestra a continuación.

	S T A T I S T I C S	
Teclado <i>Qwerty</i>	0 1 0 1 1 0 1 1 2 0	(suma = 7)
Teclado <i>Dvorak</i>	0 0 0 0 0 0 0 0 1 0	(suma = 1)

Usando este sistema de calificación con cada una de las 52 palabras en el preámbulo de la Constitución de Estados Unidos, obtenemos los valores siguientes.

Calificaciones de las palabras con un teclado Qwerty:

2 2 5 1 2 6 3 3 4 2 4 0 5
 7 7 5 6 6 8 10 7 2 2 10 5 8
 2 5 4 2 6 2 6 1 7 2 7 2 3
 8 1 5 2 5 2 14 2 2 6 3 1 7

Calificaciones de las palabras con un teclado Dvorak:

2 0 3 1 0 0 0 0 2 0 4 0 3
 4 0 3 3 1 3 5 4 2 0 5 1 4
 0 3 5 0 2 0 4 1 5 0 4 0 1
 3 0 1 0 3 0 1 2 0 0 0 1 4

- Cree una tabla de frecuencias para las calificaciones de las palabras con el teclado *Qwerty*. Utilice clases de 0-2, 3-5, 6-8, 9-11 y 12-14. Incluya una columna para la frecuencia relativa.
- Cree una tabla de frecuencias para las calificaciones de las palabras con el teclado *Dvorak*, utilice las mismas clases que en el inciso a. Incluya una columna para la frecuencia relativa.
- Con base en sus resultados de los incisos a y b, ¿cuál acomodo del teclado es más fácil para escribir? Explique.

20. Clases dobles. Los estudiantes en una clase de estadística realizan una encuesta de transporte de estudiantes en su preparatoria. Entre otros datos, para cada estudiante, ellos registran la edad y el tipo de transporte de su casa a la escuela. La tabla siguiente proporciona algunos datos que se recolectaron. Para la edad: 1 = 14 años, 2 = 15 años, 3 = 16 años, 4 = 17 años, 5 = 18 años. Para transporte: 1 = caminando, 2 = autobús escolar, 3 = transporte público, 4 = manejando, 5 = otro.

Estudiante	Edad	Transporte	Estudiante	Edad	Transporte
1	1	1	11	3	5
2	5	1	12	5	5
3	2	2	13	1	2
4	3	5	14	5	5
5	4	3	15	5	5
6	1	1	16	4	4
7	5	2	17	2	2
8	2	1	18	3	1
9	3	4	19	3	3
10	1	3	20	1	4

- Clasifique las dos variables, *edad* y *transporte*, como cualitativa o cuantitativa, y proporcione el nivel de medida de cada una.
- Para analizarse o exhibirse, los datos deben organizarse respecto a ambas variables. Cuente el número de estudiantes de cada una de las 25 categorías edad/transporte y llene las celdas vacías en la tabla siguiente.

		Transporte				
		1	2	3	4	5
Edad	1					
	2					
	3					
	4					
	5					



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 3 en www.aw.com/bbt.

- Tablas de energía.** El sitio web del Departamento de Información de la Energía de Estados Unidos ofrece muchas tablas relativas al uso de energía, precios de energía y contaminación. Explore la selección de tablas. Encuentre una tabla de datos que sea de su interés y conviértala a una tabla de frecuencias apropiada. Discuta brevemente lo que pueda aprender de la tabla de frecuencias que sea menos obvio que la tabla de datos sin procesar.
- Especies en peligro de extinción.** El sitio web del Centro de Monitoreo para la Conservación Mundial en Gran Bretaña proporciona información sobre especies de animales extintas, en peligro de extinción y amenazadas. Explore estos datos y resuma algunos de sus hallazgos más importantes con tablas de frecuencias.
- Datos.** La *relación al ombligo* se define como la altura de una persona dividida entre la altura (desde el piso) a su ombligo. Y una vieja teoría dice que, en promedio, la razón al ombligo es la razón áurea (a veces llamada razón de oro o dorada): $(1 + \sqrt{5})/2$. Mida la relación al ombligo de todas las personas de su clase. ¿Qué porcentaje tiene una razón dentro del 5% de la razón áurea? ¿Qué porcentaje de estudiantes tienen una razón áurea dentro del 10% de la razón áurea? ¿Parece confiable esta vieja teoría?

- Su propia tabla de frecuencias (sin clases).** Recolecte su propia información de frecuencias para algún conjunto de categorías que *no* requerirán clasificarse. (Por ejemplo, podría recolectar información preguntando a sus amigos para que hagan una prueba de sabor de alguna marca de galletas). Indique cómo recolectó sus datos y haga una lista de todos sus datos sin procesar. Luego resuma los datos en una tabla de frecuencia y también incluya una columna para la frecuencia acumulada, si ésta es apropiada. (La frecuencia acumulada es apropiada para todos los datos excepto para aquéllos en el nivel nominal de medida).
- Su propia tabla de frecuencias (sin clases).** Recolecte su propia información para algún conjunto de categorías que *requerirán* clasificarse (por ejemplo, pesos de sus amigos o calificaciones en algún examen reciente). Indique cómo recolectó sus datos, y haga una lista de todos sus datos sin procesar. Incluya una columna para la frecuencia relativa y para la frecuencia acumulada.



EN LAS NOTICIAS



- Tablas de frecuencias.** Encuentre un artículo reciente que incluya algún tipo de tabla de frecuencia. Describa brevemente la tabla y cómo es útil para el reporte de la noticia. ¿Considera que la tabla fue elaborada de la mejor manera posible? Si es así ¿por qué? Si no, ¿usted qué habría hecho diferente?
- Frecuencias relativas.** Encuentre un artículo reciente que proporcione al menos alguna información en forma de frecuencias relativas. Describa brevemente la información, y analice por qué las frecuencias relativas fueron útiles en este caso.
- Frecuencias acumuladas.** Encuentre una noticia reciente que proporcione al menos alguna información en la forma de frecuencia acumulada. Describa brevemente la información y analice por qué las frecuencias acumuladas fueron útiles en este caso.
- Datos de temperatura.** Busque en un periódico un reporte de clima que liste las temperaturas más altas esperadas en muchas ciudades de Estados Unidos. (Saque una fotocopia para que su profesor pueda verla). Seleccione clases apropiadas, construya una tabla de frecuencias para la información de temperaturas más altas. Incluya una columna para la frecuencia relativa y para la frecuencia acumulada. Describa brevemente cómo y por qué eligió sus clases.

3.2 Distribución gráfica de datos

Una tabla de frecuencias muestra cómo una variable se distribuye en las categorías seleccionadas, por lo que decimos que resume la **distribución** de los datos. Aunque las tablas pueden ser extraordinariamente útiles, con frecuencia obtenemos una mayor idea de la distribución con un dibujo o gráfica. En esta sección estudiaremos algunos de los métodos más comunes para mostrar distribuciones de datos.

Definición

La **distribución** de una variable se refiere a la forma en que sus valores se extienden sobre todos los valores posibles. Podemos resumir una distribución en una tabla o mostrar una distribución de manera visual por medio de una gráfica.

Gráficas de barras, diagramas de puntos y diagramas de Pareto

Una **gráfica de barras** es una de las maneras más sencillas de representar una distribución. Las gráficas de barras son utilizadas para datos cualitativos. Cada barra representa la frecuencia (o frecuencia relativa) de una categoría: cuanto mayor sea la frecuencia más larga será la barra. Las barras pueden ser verticales u horizontales.

Ahora crearemos una gráfica de barras verticales con los datos de la calificación de un ensayo de la tabla 3.1. Necesitamos cinco barras, una para cada una de las cinco categorías (las calificaciones A, B, C, D, F). La altura de cada barra debe corresponder a la frecuencia de su categoría. La figura 3.1 muestra el resultado. Observe las siguientes características clave en la construcción de la gráfica:

- Puesto que la mayor frecuencia es 9 (la correspondiente a las calificaciones C), elegimos hacer que la escala vertical vaya de 0 a 10. Esto asegura que incluso la barra más alta no toque la parte superior de la gráfica.
- La gráfica no debe ser demasiado pequeña ni demasiado grande. En este caso parece correcto seleccionar una altura total de 5 centímetros, lo que es conveniente ya que significa que cada centímetro corresponde a una frecuencia de 2.

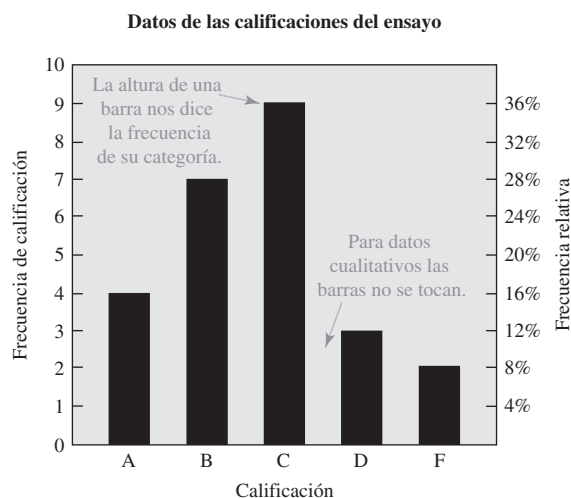


Tabla 3.1 (repetida)	
Calificación	Frecuencia
A	4
B	7
C	9
D	3
F	2
Total	25

Figura 3.1 Gráfica de barras para las calificaciones del ensayo de la tabla 3.1.

- La altura de cada barra debe ser proporcional a su frecuencia. Por ejemplo, como cada centímetro corresponde a una frecuencia de 2, la barra que representa una frecuencia de 4 debe tener una altura de 2 centímetros.
- Ya que los datos son cualitativos, la anchura de las barras no tiene un significado especial, y no hay razón para que se toquen entre ellas. Por tanto, dibújelas con anchuras uniformes.

La rotulación de las gráficas es muy importante. Sin etiquetas apropiadas una gráfica carece de sentido. El resumen siguiente proporciona las etiquetas importantes para casi cualquier gráfica. Observe cómo estas reglas se aplicaron en la figura 3.1.

Etiquetas importantes para las gráficas

Título/pie de figura: la gráfica debe tener un título o un pie de figura (o ambos) que explique lo que se muestra y, si es aplicable, liste la fuente de los datos.

Escala y etiqueta del eje vertical: los números a lo largo del eje vertical deben indicar claramente la escala. Los números deben alinearse con las *marcas*; las marcas a lo largo del eje ubican de manera precisa los valores numéricos. Incluya una etiqueta que describa la variable que se muestra en el eje vertical.

Escala y etiqueta del eje horizontal: las categorías deben indicarse con claridad a lo largo del eje horizontal. (Las marcas pueden no ser necesarias para datos cualitativos, pero sí lo son para datos cuantitativos). Incluya una etiqueta que describa la variable que se muestra en el eje horizontal.

Leyenda: si se exhiben varios conjuntos de datos en una sola gráfica, incluya una leyenda o clave para identificar los conjuntos individuales de datos.

Un **diagrama de puntos** es una variación de una gráfica de barras en la que utilizamos puntos en lugar de barras para representar las frecuencias. Cada punto representa un valor de dato; por ejemplo, una pila con 4 puntos significa una frecuencia de 4. La figura 3.2 muestra un diagrama de puntos para el conjunto de datos del ensayo. Los diagramas de puntos son convenientes cuando se hacen gráficas de datos sin procesar, ya que puede contar los datos haciendo un punto por cada dato. Luego puede decidir convertir la gráfica en una gráfica de barras para un reporte formal.

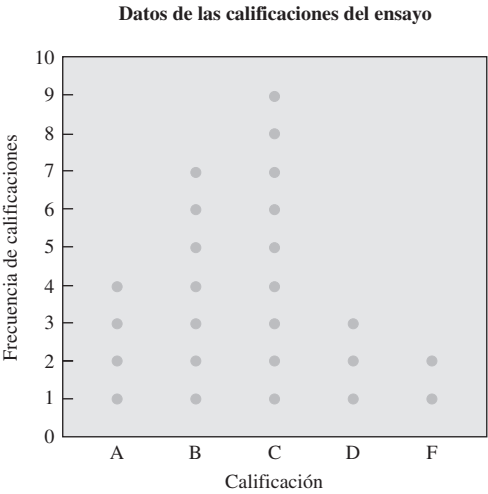


Figura 3.2 Gráfica de puntos para los datos del ensayo de la tabla 3.1.

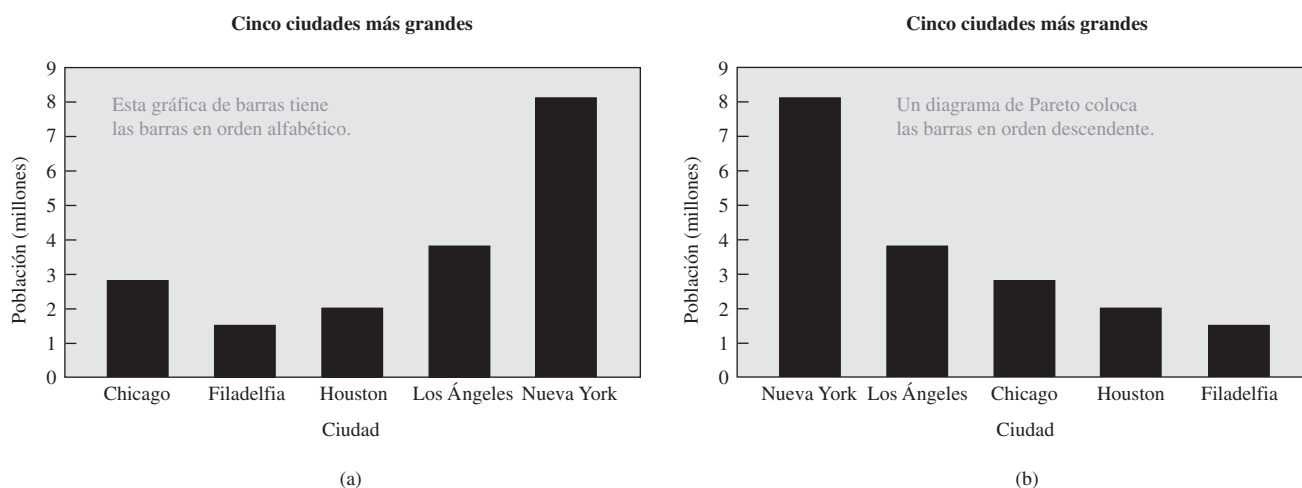


Figura 3.3 (a) Gráfica de barras que muestra las poblaciones para las cinco ciudades más grandes de Estados Unidos (2005). (b) Diagrama de Pareto para los mismos datos. *Fuente:* Oficina del Censo de Estados Unidos.

La figura 3.3a muestra una gráfica de barras de poblaciones para las cinco ciudades más grandes en Estados Unidos. En este caso, las cinco ciudades son las categorías y sus poblaciones son las frecuencias; las ciudades se listan en orden alfabético. La figura 3.3b muestra los mismos datos, pero con las barras acomodadas en orden descendente. Una gráfica de barras en la cual éstas se acomodan en orden de su frecuencia se denomina **diagrama de Pareto**. Observe que el reacomodo de las barras sólo tiene sentido si los datos son cualitativos en el nivel nominal de medida, como son cuando las categorías son ciudades. Por ejemplo, no tendría sentido hacer un diagrama de Pareto para los datos de las calificaciones de nivel ordinal de la figura 3.1, ya que colocar las barras en orden de frecuencias pondría las calificaciones en el orden C, B, A, D, F.

UN MOMENTO DE REFLEXIÓN

¿Sería práctico construir un diagrama de puntos para la información de la población en la figura 3.3? ¿Tendría sentido hacer un diagrama de Pareto para los datos concernientes a calificaciones del SAT? Explique.

Definiciones

Una **gráfica de barras** consiste en barras que representan las frecuencias (o frecuencias relativas) para categorías particulares. Las longitudes de las barras son proporcionales a las frecuencias.

Un **diagrama de puntos** es similar a una gráfica de barras, salvo que cada dato individual se representa por un punto.

Un **diagrama de Pareto** es una gráfica con barras acomodadas en orden de frecuencias. Los diagramas de Pareto tienen sentido sólo para datos en el nivel nominal de medida.

A propósito...

Los diagramas de Pareto los inventó el economista italiano Vilfredo Pareto (1848-1923). Pareto es mejor conocido por el desarrollo de métodos de análisis de distribuciones de ingresos, pero quizá sus contribuciones más importantes fueron en el desarrollo de nuevas formas de aplicación de la matemática o la estadística al análisis económico.

EJEMPLO 1 Emisiones de dióxido de carbono

El dióxido de carbono se libera a la atmósfera principalmente por la combustión de combustibles fósiles (petróleo, carbón, gas natural). La tabla 3.11 lista los ocho países que anualmente emiten la mayoría del dióxido de carbono. Construya diagramas de Pareto para las emisiones totales y para el promedio de emisiones por persona. Analice por qué los dos diagramas parecen tan diferentes.

NOTA TÉCNICA

La tabla 3.11 da las emisiones en términos del peso del carbono contenido en el dióxido de carbono emitido (CO₂), pero no incluye el peso del oxígeno en este último. Desenlace

Tabla 3.11 Los ocho líderes mundiales en emisión de dióxido de carbono		
País/región	Emisiones totales de dióxido de carbono (millones de toneladas métricas de carbono)	Emisiones de dióxido de carbono por persona (toneladas métricas de carbono)
Estados Unidos	1582	5.4
China	966	0.7
Rusia	438	3.0
Japón	329	2.6
India	280	0.3
Alemania	230	2.8
Canadá	164	5.2
Reino Unido	154	2.6

Fuente: Departamento de Energía de Estados Unidos. Datos publicados en 2007, que muestran las emisiones de 2003.

Solución Las categorías son los países y las frecuencias los valores. Puesto que el rango de valores para las emisiones totales de dióxido de carbono va de 154 a 1 582, un rango de 0 a 1 600 es una buena elección para la escala vertical. Para construir un diagrama de Pareto ponemos las barras en orden descendente de tamaño. La altura de cada barra corresponde a su valor y rotulamos la categoría (país) bajo la barra. La figura 3.4a muestra la gráfica resultante para las emisiones totales de dióxido de carbono. La figura 3.4b muestra el diagrama de Pareto para las emisiones por persona, un rango vertical de 0 a 6 abarca todos los datos.

Observe que los dos diagramas de Pareto tienen a los países en orden diferente. Esto nos indica que los mayores emisores de dióxido de carbono no necesariamente son los mayores emisores por persona. Por ejemplo, China se clasifica como el segundo mayor emisor total, pero las emisiones por persona están muy por debajo de Estados Unidos y de otras naciones desarrolladas.

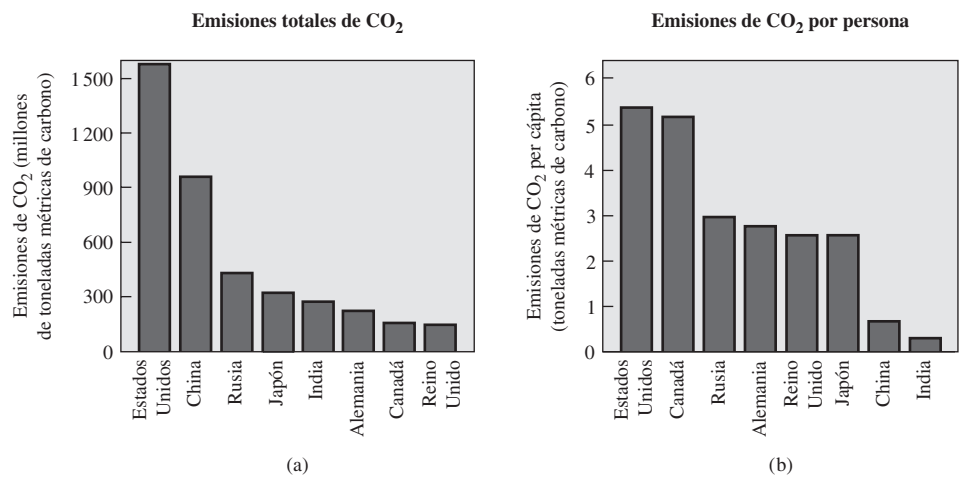


Figura 3.4 Diagramas de Pareto para (a) emisiones de dióxido de carbono por país, y (b) emisiones de dióxido de carbono por persona.

UN MOMENTO DE REFLEXIÓN

La población combinada de China e India es más de ocho veces la población de Estados Unidos, sin embargo, las emisiones de dióxido de carbono de Estados Unidos son mayores que las de China e India juntas. ¿Por qué cree que es el caso? ¿Qué consecuencias podría haber para el mundo si China e India tuviesen las mismas emisiones de dióxido de carbono por persona que Estados Unidos?

Gráficas circulares

Las **gráficas circulares** (a veces llamadas gráficas de pastel) se utilizan para mostrar distribuciones de frecuencias relativas. Una gráfica circular representa la frecuencia relativa total de 100%, y los tamaños de los sectores, o cuñas, representan las frecuencias relativas de las diferentes categorías. Las gráficas circulares se utilizan casi exclusivamente para datos cualitativos.

Como un ejemplo sencillo, considere los votantes registrados en el condado de Rochester: 25% son demócratas, 25% son republicanos y 50% son independientes. Podemos mostrar estas afiliaciones en una gráfica circular. Ya que los demócratas y los republicanos, cada uno, representan 25% de los votantes, los sectores para cada uno de ellos ocupa 25%, o un cuarto, del círculo. Los independientes representan la mitad de los votantes, por lo que su sector ocupa la mitad restante del círculo. La figura 3.5 muestra el resultado. Como siempre, observe la importancia de poner un título claro.

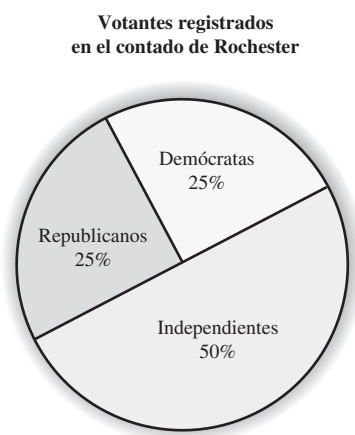


Figura 3.5 Afiliaciones partidistas de votantes registrados en el condado de Rochester.

Definición

Una **gráfica circular** es un círculo dividido de forma que cada sector representa la *frecuencia relativa* de una categoría particular. El tamaño del sector es proporcional a la frecuencia relativa. El círculo completo representa la frecuencia relativa total de 100%.

Una gráfica circular como la figura 3.5 es fácil de crear, ya que el tamaño de los sectores representa fracciones muy sencillas. Para gráficas circulares más complejas, debe tener cuidado en la medición de los ángulos o utilizar un programa que haga las medidas. Y aunque las gráficas circulares pueden ser útiles para conjuntos de datos simples, el ejemplo siguiente muestra que gráficas circulares no siempre pueden ser la mejor manera de presentar los datos.

EJEMPLO 2 Especialidad de estudiantes

La figura 3.6 es una gráfica circular que muestra las principales áreas para estudiantes de primer año de universidad. Construya un diagrama de Pareto que muestre los mismos datos. ¿Cuáles son las tres áreas de especialización más comunes? Comente sobre la facilidad relativa con la cual esta pregunta puede responderse con la gráfica circular y con el diagrama de Pareto.

A propósito...

Nacionalmente, en las elecciones de 2006, 37% de los estadounidenses en edad de votar se identificaron como demócratas, 31% como republicanos y 32% como independientes o de otros partidos. (*Resumen Estadístico de Estados Unidos*).





Figura 3.6 Principales áreas planeadas por estudiantes universitarios de primer año.

Fuente: *The Chronicle of Higher Education*.

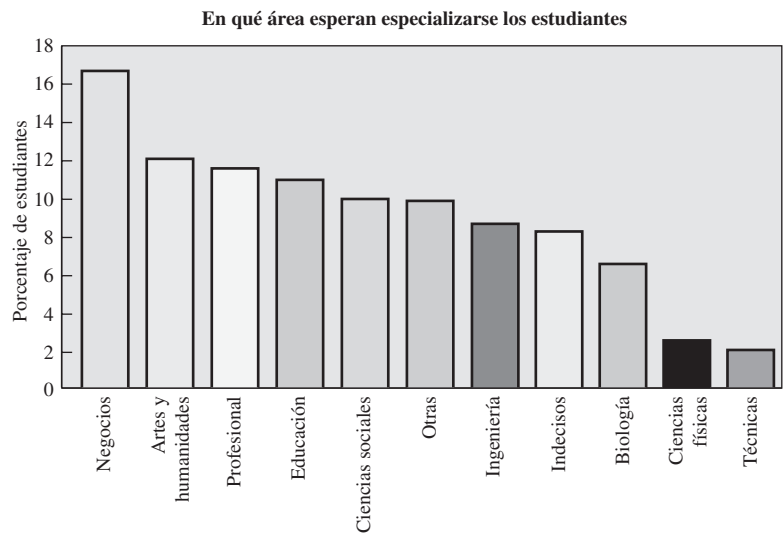


Figura 3.7 Diagrama de Pareto para los datos en la figura 3.6.

Solución La figura 3.7 muestra el diagrama de Pareto para los datos. Este diagrama hace muy obvio que las tres áreas de especialización más comunes son negocios (16.7%), artes y humanidades (12.1%) y profesional (11.6%). (“Profesional” incluye especialidades en los campos con licenciatura, tal como arquitectura, enfermería, y farmacología). En contraste, toma un poco de tiempo el estudio de la gráfica circular antes de que podamos listar las tres áreas más comunes de especialización.

UN MOMENTO DE REFLEXIÓN

El ejemplo 2 estudia una ventaja del diagrama de Pareto sobre una gráfica circular mostrando los datos relativos a áreas de especialización. ¿Considera que la gráfica circular tiene algunas ventajas sobre el diagrama de Pareto? Si es así, ¿cuáles son?

NOTA TÉCNICA

Diversos textos definen los términos *histograma* y *gráfica de barras* de forma diferente, y no existen definiciones universalmente aceptadas. En este libro una gráfica de barras es cualquier gráfica que utilice barras, y los histogramas son los tipos de barras usadas para datos cuantitativos en el nivel de intervalo o de razón de medida.

Histogramas y diagramas de tallos y hojas

La figura 3.8 muestra una gráfica de barras para la información del uso de la energía de la tabla 3.4. El eje horizontal está marcado con las categorías del uso de energía por persona por año (en millones de BTU), y el eje vertical está marcado con las frecuencias (número de estados) que corresponden a cada categoría. Como en toda gráfica de barras, las longitudes de las barras son proporcionales a las frecuencias. Sin embargo, a diferencia de las barras en las gráficas que hicimos anteriormente con datos cualitativos, las barras en esta gráfica caen en un orden natural con base en los valores de la categoría. Además, los anchos de las barras en la figura 3.8 tienen un significado específico, en este caso nos dicen el rango de valores en cada una de las categorías de uso de la energía. Este tipo de gráfica de barras, en el que las barras tienen un orden natural y las anchuras de las barras tienen un significado específico, se denomina **histograma**. Las barras en un histograma se tocan ya que no hay separación entre las categorías.

El histograma del uso de energía revela con claridad las tendencias listadas en una tabla de frecuencias como la 3.4. Sin embargo, ni la tabla de frecuencias ni el histograma proporcionan todos los detalles del conjunto original de datos de la tabla 3.3. Por ejemplo, el histograma nos dice que un estado tiene uso de energía en la categoría de 800-899 millones de BTU por persona, pero no nos dice cuál es o el valor preciso del uso de energía para ese

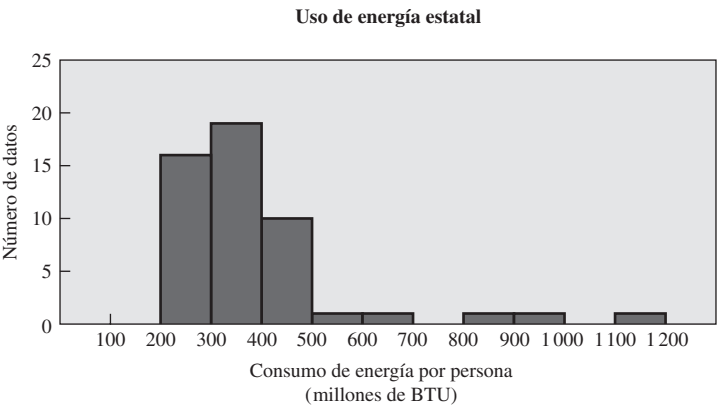


Figura 3.8 Histograma para los datos en la tabla 3.4.

estado. El **diagrama de tallos y hojas** (o *diagrama de tallos*) de la figura 3.9 proporciona una vista más detallada de los datos. Se parece a un histograma girado excepto que en lugar de barras vemos una lista de datos —en este caso, los estados— para cada categoría. Por ejemplo, ahora podemos ver que el estado en la categoría 800-899 millones de BTU por persona es Louisiana.

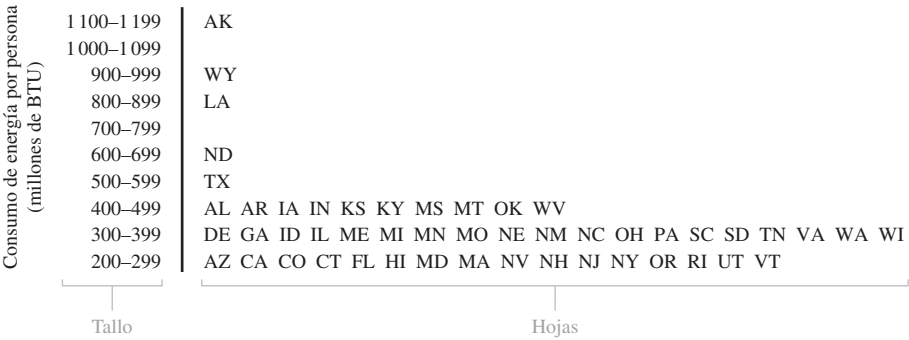


Figura 3.9 Diagrama de tallos y hojas para los datos de uso de energía de la tabla 3.3.

Otro tipo de diagrama de tallos y hojas lista los *valores* individuales. Por ejemplo, la figura 3.10 muestra un diagrama de tallos y hojas para las emisiones por persona de dióxido de carbono de la tabla 3.11. En este caso cada valor se representa con un tallo que consiste en el primer dígito (y el punto decimal) y una hoja consiste en el dígito restante. Podemos leer los valores directamente de este diagrama. Por ejemplo, la primera fila muestra los valores 0.3 y 0.7.

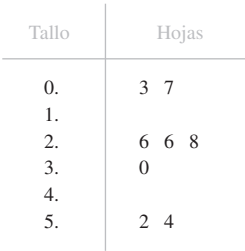


Figura 3.10 Diagrama de tallos y hojas que muestra datos numéricos, en este caso las emisiones de dióxido de carbono por persona de la tabla 3.11.

Definiciones

Un **histograma** es una gráfica de barras que muestra una distribución para datos cuantitativos (en el nivel de intervalo o de razón de medida); las barras tienen un orden natural y las anchuras de las barras tienen un significado específico.

Un **diagrama de tallos y hojas** (o *diagrama de tallos*) es muy parecido a un histograma girado, salvo que en lugar de barras vemos un listado de datos.

Tabla 3.12 Edades de actrices en el momento de recibir el premio de la academia 1970-2007

Edad	Número de actrices
20-29	8
30-39	18
40-49	8
50-59	0
60-69	2
70-79	1
80-89	1

EJEMPLO 3 Actrices ganadoras de un Oscar

La tabla 3.12 muestra las edades (en el momento que ganaron el premio) de actrices ganadoras del premio de la academia desde 1970 a 2007. Construya un histograma para mostrar estos datos. Analice los resultados.

Solución La frecuencia más alta es 18 (actrices), por lo que elegimos 20 como la altura del histograma. Como en cualquier gráfica de barras, la altura de cada barra corresponde a la frecuencia para la categoría. La anchura de cada barra cubre 10 años completos, por lo que las barras se tocan. La figura 3.11 muestra el resultado. Vemos que es más probable que las actrices ganen Oscars cuando son muy jóvenes. En contraste, los actores hombres tienden a ganar Oscars a edades más grandes (vea el ejercicio 13 en la sección 3.1). Muchas actrices creen que estos hechos reflejan una forma sutil de discriminación: los productores de Hollywood rara vez producen películas que caractericen a mujeres mayores en papeles principales.

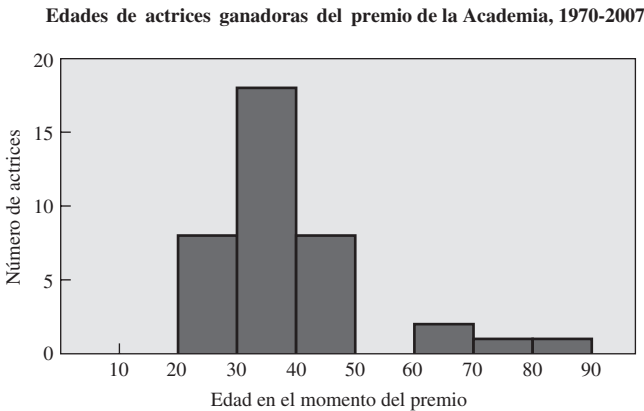


Figura 3.11 Histograma para edades de actrices ganadoras del Oscar.

Apropósito. . .

La actriz de mayor edad ganadora de un Oscar fue Jessica Tandy, quien ganó en 1990 por la película *Driving Miss Daisy*.

UN MOMENTO DE REFLEXIÓN

¿Qué información adicional necesitaría para crear un diagrama de tallos y hojas para las edades de actrices cuando ganaron el premio de la academia? ¿Cómo se vería el diagrama de tallos y hojas?

NOTA TÉCNICA

Una gráfica de líneas creada con base en un conjunto de datos de frecuencia a veces se conoce como *polígono de frecuencias*, ya que consiste en varios segmentos de rectas que toman la forma de una figura de muchos lados.

Gráficas de líneas

Al igual que un histograma, una **gráfica de líneas** muestra una distribución de datos cuantitativos. Sin embargo, en lugar de usar barras, un diagrama de líneas conecta una serie de puntos. La posición vertical de cada punto representa una frecuencia; los puntos van donde iría la parte superior de una barra en un histograma. La figura 3.12 muestra una gráfica de líneas para los datos de energía de la tabla 3.4. Para comparar sobreponemos el histograma de la figura 3.8.

Observe una sutileza importante al interpretar una gráfica de líneas. Las posiciones horizontales de los puntos corresponden a los *centros* de las clases. Por ejemplo, el punto que representa la clase para 300-399 millones de BTU está ubicado en 349.5 millones de BTU. Por tanto, si no sabe algo más, podría considerar que el punto significa que 19 estados tienen un uso de la energía de *exactamente* 349.5 millones de BTU por persona, cuando en realidad 19 estados están en el *rango* de 300-399 millones de BTU por persona.

Definición

Un **diagrama o gráfica de líneas** muestra una distribución de datos cuantitativos como una serie de puntos conectados por medio de rectas. Para cada punto, la posición horizontal es el *centro* de la clase que representa y la posición vertical es el valor de la frecuencia para esa clase.

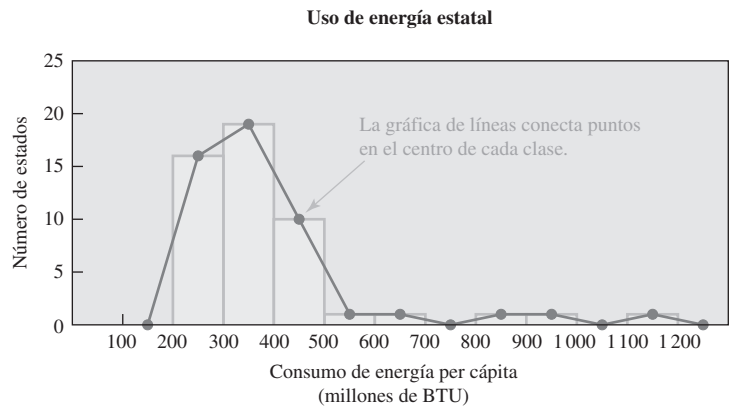


Figura 3.12 Gráfica de líneas para los datos de uso de energía, con un histograma superpuesto por comparación.

EJEMPLO 4 Distribución de edades

La tabla 3.13 muestra la distribución de la población de Estados Unidos de acuerdo con categorías de edades. Construya un histograma y una gráfica de líneas para estos datos. Si hiciera una gráfica similar para edades, dentro de 20 años ¿cómo esperaría que fuese diferente?

Tabla 3.13 Distribución estimada de edades de la población de Estados Unidos, 2007					
Categoría de edad	Población (millones)	Categoría de edad	Población (millones)	Categoría de edad	Población (millones)
0-4	21.0	35-39	19.0	70-74	9.0
5-9	19.4	40-44	20.4	75-79	7.2
10-14	19.9	45-49	22.2	80-84	5.6
15-19	21.7	50-54	22.0	85-89	3.5
20-24	21.2	55-59	19.2	90-94	1.6
25-29	19.8	60-64	16.3	95-99	0.6
30-34	19.0	65-69	12.2	100 o más	0.1

Fuente: Oficina del Censo de Estados Unidos.

Solución Los datos ya están separados en clases con intervalos de 5 años (excepto por la última categoría de “100 o más”). Construimos el histograma representando la frecuencia (población) en cada categoría con una barra. La figura 3.13 muestra el resultado. Sobre este histograma está sobrepuesta una gráfica de líneas en la que un punto se colocó en el centro de cada categoría. Por ejemplo, el punto que representa la clase de 25 a 29 años de edad está ubicado en 27.5 sobre el eje horizontal. Por consistencia, mostramos la categoría “100 o más” como si fuese 100-104, pero la etiqueta en el eje horizontal muestra lo que realmente significa. Muchas tablas de frecuencia incluyen categorías con un extremo abierto, como “100 o más” cuando se tiene una pequeña cantidad de valores extremos en un rango amplio.

Si hiciésemos una gráfica similar dentro de 20 años, esperaríamos ver dos diferencias principales. Primera, todas las frecuencias serían mayores debido al crecimiento poblacional. Segunda, habría un mayor crecimiento en las edades superiores que en las edades inferiores, consecuencia de los avances en tecnología médica que permiten a más gente vivir más años.

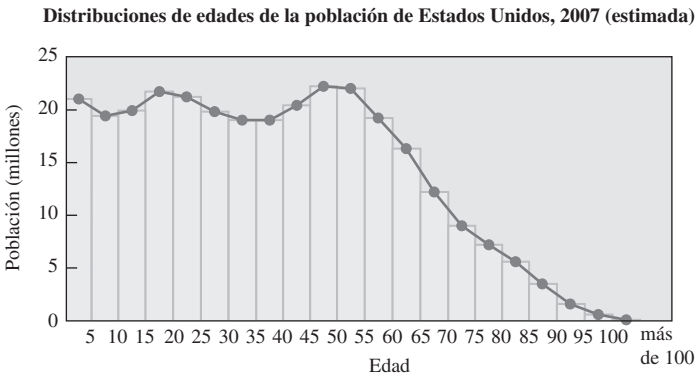


Figura 3.13 Histograma y gráfica de líneas para los datos de edades en la tabla 3.13.

Diagramas de series de tiempo

La tabla 3.14 muestra cómo la tasa de homicidios en Estados Unidos ha cambiado a través del tiempo. Las categorías son los años y los datos son las tasas de homicidios (medidas en asesinatos por cada 100 000 personas). Podemos representar estos datos, ya sea con un histograma o con una gráfica de líneas; la figura 3.14 muestra una gráfica de líneas. Puesto que en este caso el eje horizontal representa tiempo, decimos que esta gráfica es una **gráfica de serie de tiempo**.

Tabla 3.14 Tasa de homicidios en Estados Unidos por cada 100 000 personas, desde 1960					
Año	Homicidios (por 100 000 personas)	Año	Homicidios (por 100 000 personas)	Año	Homicidios (por 100 000 personas)
1960	5.1	1976	8.8	1992	9.3
1961	4.8	1977	8.8	1993	9.5
1962	4.6	1978	9.0	1994	9.0
1963	4.6	1979	9.7	1995	8.2
1964	4.9	1980	10.2	1996	7.4
1965	5.1	1981	9.8	1997	6.8
1966	5.6	1982	9.1	1998	6.3
1967	6.2	1983	8.3	1999	5.7
1968	6.9	1984	7.9	2000	5.5
1969	7.3	1985	7.9	2001	5.6
1970	7.9	1986	8.6	2002	5.6
1971	8.6	1987	8.3	2003	5.7
1972	9.0	1988	8.4	2004	5.5
1973	9.4	1989	8.7	2005	5.8
1974	9.8	1990	9.4		
1975	9.6	1991	9.8		

Fuente: Oficina de Estadísticas de Justicia de Estados Unidos.

Definición

Un histograma o una gráfica de líneas en la que el eje horizontal representa *tiempo* se denomina **gráfica de series de tiempo**.



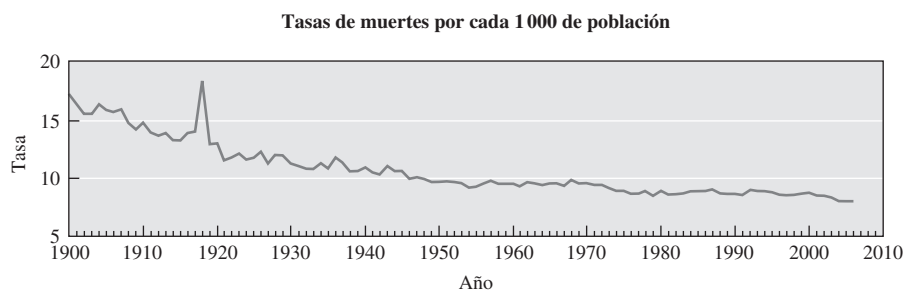
Figura 3.14 Gráfica de serie de tiempo para los datos de tasa de homicidios de la tabla 3.14.

UN MOMENTO DE REFLEXIÓN

Busque información ampliada de la tabla 3.14 y la figura 3.14 posterior a 2005. ¿Observa alguna tendencia en la tasa de homicidios?

EJEMPLO 5 Disminución de tasa de mortalidad

La figura 3.15 muestra una gráfica de serie de tiempo para la tasa de mortalidad (muertes por 1000 personas) en Estados Unidos desde 1900. (Por ejemplo, la tasa de mortalidad en 1905 de 15 significa que, por cada 1000 personas que vivían al inicio de 1905, 15 de ellas morían durante el año). Analice la tendencia general y las razones para la tendencia. Además, considere el pico en 1919: si alguien le dijese que este pico se debió a las muertes en batalla en la Primera Guerra Mundial, ¿lo creería? Explique.



Fuente: National Center for Health Statistics

Figura 3.15 Tasas históricas de muertes en Estados Unidos por cada 1000 personas.

Solución La tendencia general en las tasas de mortalidad claramente es descendente, presumiblemente a consecuencia de los adelantos en la ciencia médica. Por ejemplo, las enfermedades bacteriales, como la neumonía, fueron la principal causa de mortandad a principios de 1900, pero son muy curables con los antibióticos actuales. El pico en 1919 *sí* coincide con el final de la Primera Guerra Mundial. Sin embargo, si el pico fuera debido a las bajas en batalla, podríamos esperar que se extendiera durante los años de la Primera Guerra Mundial, además de ver un pico similar durante la Segunda Guerra Mundial. La ausencia de estas características sugiere que el pico de 1919 podría no deberse a la Primera Guerra Mundial. De hecho, la razón del pico fue una epidemia mortal de influenza.

A propósito...

La epidemia de influenza de 1919 mató a 850 000 personas en Estados Unidos y un estimado de 20 millones en todo el mundo. La mayoría de las muertes fueron causadas por infecciones bacteriales secundarias que serían curables actualmente con antibióticos.

Sección 3.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Distribución.** ¿Qué queremos decir por la distribución de datos? ¿Cómo es más útil un histograma que una lista de valores muestrales para la comprensión de la distribución?
- Conjuntos pequeños de datos.** Si un conjunto de datos sólo tiene cinco valores, ¿por qué no deberíamos preocuparnos por construir un histograma?
- Gráficas de series de tiempo.** ¿Qué tipo de datos se requieren para la construcción de una gráfica de series de tiempo? y ¿qué revela de los datos una gráfica de series de tiempo?
- Histograma y diagrama de tallos y hojas.** Suponga que un conjunto de datos se utiliza para construir un histograma y un diagrama de tallos y hojas. Usando sólo el histograma, ¿es posible recrear la lista de datos originales? Usando sólo el diagrama de tallos y hojas, ¿es posible recrear la lista de datos originales? ¿Cuál es una ventaja de un diagrama de tallos y hojas sobre un histograma?

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Histograma.** Se utilizó un histograma para describir una muestra de alturas clasificadas de acuerdo con el color de los ojos del sujeto.
- Diagrama de serie de tiempo.** Un diagrama de serie de tiempo se usaría para mostrar la utilidad neta de diferentes fabricantes de computadoras en 2007.
- Diagrama de Pareto.** Una ingeniera de control de calidad quiere llevar la atención hacia las causas principales de defectos, por lo que ella utiliza un diagrama de Pareto para ilustrar las frecuencias de las diferentes causas de defectos.
- Gráfica circular.** Una desventaja importante de las gráficas circulares es que carecen de una escala apropiada.

Conceptos y aplicaciones

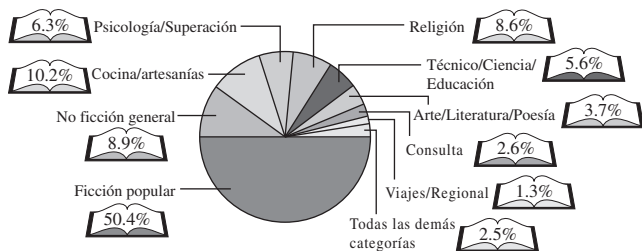
Representación gráfica más apropiada. Los ejercicios 9 al 12 describen conjuntos de datos, pero no dan los datos reales. Para cada conjunto de datos indique el tipo de gráfica que considera más adecuado para mostrar los datos, si estuvieran disponibles. Explique su elección.

- Calificaciones del SAT.** Calificaciones del SAT de estudiantes de estadística.
- Color de ojos.** Colores de ojos de actrices.
- Causas de muertes.** Causas de muertes de personas que fallecieron el año pasado.

- Súper Tazón.** Costo de comerciales televisivos durante el medio tiempo del juego del Súper Tazón para cada año, iniciando en el primer Súper Tazón en 1967.

- Qué leen las personas.** La gráfica circular en la figura 3.16 muestra los resultados de una encuesta acerca de lo que están leyendo las personas.

- Resuma estos datos en una tabla de frecuencias relativas.
- Construya un diagrama de Pareto para estos datos.
- ¿Cuál cree que es la mejor representación de los datos: la gráfica circular o el diagrama de Pareto? ¿Por qué?



Fuente: Book Industry Study Group

Figura 3.16 Fuente: Wall Street Journal Almanac.

- Histograma.** El histograma generado con STATDISK en la figura 3.17 describe los niveles de cotinina (en miligramos por mililitro) de una muestra de sujetos que fuman cigarrillos. La cotinina es un metabólico de la nicotina, lo que significa que la cotinina es producida por el cuerpo cuando la nicotina es absorbida. Los datos provienen de la Tercera Encuesta Nacional de Salud y Nutrición.

- ¿Cuántos sujetos están representados en el histograma?
- ¿Cuántos sujetos tienen niveles de cotinina inferiores a 400?
- ¿Cuántos sujetos tienen niveles de cotinina por arriba de 150?
- ¿Cuál es el nivel más alto de cotinina de un sujeto representado en este histograma?

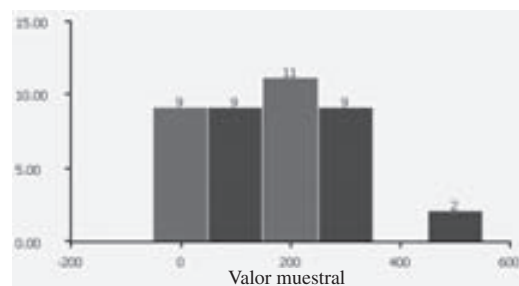


Figura 3.17

- 15. Pesos de Coca.** El ejercicio 11 en la sección 3.1 requirió la construcción de una tabla de frecuencias de los pesos (en libras) de 36 latas de Coca regular. Utilice esa tabla de frecuencias para construir el histograma correspondiente.
- 16. Pesos de Coca de dieta.** El ejercicio 12 en la sección 3.1 requirió la construcción de una tabla de frecuencias de los pesos (en libras) de 36 latas de Coca de dieta. Utilice esa tabla de frecuencias para construir el histograma correspondiente.
- 17. Actores ganadores de un Oscar.** El ejercicio 13 en la sección 3.1 requirió la construcción de una tabla de frecuencias de las edades de actores hombres ganadores del premio de la academia en el momento que ganaron su premio. Utilice esa tabla de frecuencias para construir el histograma correspondiente.
- 18. Temperaturas corporales.** El ejercicio 14 en la sección 3.1 requirió la construcción de una tabla de frecuencias de una lista de temperaturas corporales (°F) de sujetos elegidos aleatoriamente. Utilice esa tabla de frecuencias para construir el histograma correspondiente.
- 19. Búsqueda de trabajo.** Se llevó a cabo un estudio para determinar cómo obtenía trabajo la gente. La tabla siguiente lista los datos de 400 sujetos elegidos aleatoriamente. Los datos están basados en resultados del Centro Nacional de Estrategias Profesionales. Construya un diagrama de Pareto que corresponda a los datos dados. Si alguien pudiera obtener un trabajo, ¿cuál parece ser el enfoque más apropiado?

Fuentes de trabajo de los encuestados	Frecuencia
Anuncios	56
Compañías de búsqueda de ejecutivos	44
Contactos profesionales	280
Envío masivo de correo	20

- 20. Fuentes de trabajo.** Consulte los datos del ejercicio 19 y construya una gráfica circular. Compare la gráfica circular con el diagrama de Pareto. ¿Puede determinar cuál gráfica es más efectiva en mostrar la importancia relativa de las fuentes de trabajo?
- 21. Causas de descarrilamiento de trenes.** Un análisis de los incidentes de descarrilamiento de trenes mostró que 23 descarrilamientos fueron causados por vías en mal estado, 9 debidos a fallas en el equipo, 12 atribuidas a error humano y 6 a otras causas (con base en información de la Administración Federal de Ferrocarriles). Construya una gráfica circular que represente los datos dados.
- 22. Análisis de las causas de descarrilamiento de trenes.** Consulte la información dada en el ejercicio 21 y construya un diagrama de Pareto. Compare el diagrama de Pareto con la gráfica circular. ¿Puede determinar cuál gráfica es más efectiva en mostrar la importancia relativa de las causas de descarrilamiento de trenes?

- 23. Gráfica de puntos.** Consulte la información del teclado QWERTY en el ejercicio 19 de la sección 3.1 y construya una gráfica de puntos.

- 24. Gráfica de puntos.** Consulte la información del teclado Dvorak en el ejercicio 19 de la sección 3.1 y construya una gráfica de puntos. Compare el resultado con la gráfica de puntos del ejercicio 23. Con base en los resultados, ¿qué configuración de teclado parece ser mejor? Explique.

- 25. Serie de tiempo de DJIA.** Los valores siguientes son los valores altos del índice industrial Dow Jones (DJIA, por sus siglas en inglés) anuales iniciando con 1980. Los datos están acomodados en orden por renglones. Construya una gráfica de serie de tiempo y luego determine si parece haber una tendencia. ¿Un inversionista podría sacar provecho de esta tendencia?

1 000	1 024	1 071	1 287	1 287	1 553
1 956	2 722	2 184	2 791	3 000	3 169
3 413	3 794	3 978	5 216	6 561	8 259
9 374	11 568	11 401	11 350	10 635	10 454
10 855	10 941	12 464			

- 26. Serie de tiempo para muertes por vehículos de motor.** Los valores siguientes son cifras de muertes causadas por vehículos de motor en Estados Unidos por años, iniciando con 1980. Los datos están acomodados por renglones. Construya una gráfica de serie de tiempo y luego determine si parece haber una tendencia. Si es así, proporcione una posible explicación.

51 091	49 301	43 945	42 589	44 257	43 825
46 087	46 390	47 087	45 582	44 599	41 508
39 250	40 150	40 716	41 817	42 065	42 013
41 501	41 717	41 945	42 196	43 005	42 884
42 836	43 443				

- 27. Diagrama de tallos y hojas.** Construya un diagrama de tallos y hojas de estas calificaciones de examen: 67, 72, 85, 75, 89, 89, 88, 90, 99, 100. ¿Cómo muestra el diagrama de tallos y hojas la distribución de los datos?

- 28. Diagrama de tallos y hojas.** A continuación se listan las duraciones (en minutos) de películas animadas para niños. Construya un diagrama de tallos y hojas. ¿El diagrama de tallos y hojas muestra la distribución de los datos? Si es así, ¿cómo?

83	88	120	64	69	71	76	74	75	76	75
75	79	80	78	78	83	77	71	83	80	73
72	82	74	84	90	89	81	81	90	79	92
82	89	82	74	86	76	81	75	75	77	70
75	64	73	74	71	94					



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 3 en www.aw.com/bbt.

- 29. Emisiones de CO₂.** En el sitio web de *International Energy Annual*, publicado por el Departamento de Información de

Energía (EIA, por sus siglas en inglés) busque información actualizada de emisiones internacionales de dióxido de carbono. Cree una gráfica que incluya la última información. Analice cualquier información importante de su gráfica.

- 30. Tabla de energía.** Explore el conjunto completo de tablas de energía en el sitio web del EIA de Estados Unidos. Seleccione una tabla que le parezca interesante y construya una gráfica de sus datos. Puede elegir cualquiera de los tipos de gráficas analizados en esta sección. Explique cómo hizo su gráfica y discuta brevemente qué puede aprender de ella.

- 31. Resumen estadístico.** Vaya al sitio web del *Statistical Abstract of the United States*. Explore la selección de “tablas frecuentemente consultadas”. Seleccione una tabla que le interese y construya una gráfica de sus datos. Puede elegir cualquiera de los tipos de gráficas analizados en esta sección. Explique cómo hizo su gráfica y discuta brevemente qué puede aprender de ella.

- 32. Datos del ombligo.** Cree una exhibición gráfica apropiada de los datos del ombligo recolectados en el ejercicio 23 de la sección 3.1. Analice cualesquiera propiedades especiales de esta distribución.

EN LAS NOTICIAS

- 33. Gráficas de barras.** Encuentre una noticia reciente que incluya una gráfica de barras con categorías cualitativas de datos.

- Explique brevemente qué muestra la gráfica de barras y analice si ayuda al punto que sostiene la noticia. ¿Las etiquetas son claras?
- Analice brevemente si la gráfica de barras podría adaptarse como una gráfica de puntos.
- ¿La gráfica ya es un diagrama de Pareto? Si es así, explique por qué considera que fue mostrada de esta manera. Si no, ¿considera que sería más clara si las barras fuesen acomodadas para construir un diagrama de Pareto? Explique.

- 34. Gráficas circulares.** Encuentre una noticia reciente que incluya una gráfica circular. Analice brevemente

la efectividad de la gráfica circular. Por ejemplo, ¿sería mejor si los datos fueran presentados en una gráfica de barras en lugar de una gráfica circular? ¿La gráfica circular podría mejorar de alguna otra forma?

- 35. Histogramas.** Encuentre una noticia reciente que incluya un histograma. Explique brevemente qué muestra el histograma y analice si ayuda al punto que sostiene la noticia. ¿Las etiquetas son claras? ¿El histograma es una gráfica de serie de tiempo? Explique.

- 36. Gráficas de líneas.** Encuentre una noticia reciente que incluya una gráfica de líneas. Explique brevemente qué muestra la gráfica de líneas y analice si ayuda al punto que sostiene la noticia. ¿Las etiquetas son claras? ¿La gráfica de líneas es una gráfica de serie de tiempo? Explique.

3.3 Gráficas en los medios

Las gráficas básicas que hemos estudiado hasta aquí son sólo el inicio de las muchas formas de describir los datos de manera visual. En esta sección exploraremos algunos de los tipos más complejos de gráficas que son comunes en los medios.

Gráfica de barras múltiple y gráficas de líneas

Una **gráfica de barras múltiple** es una extensión sencilla de una gráfica de barras común: tiene dos o más conjuntos de barras que permiten la comparación entre dos o más conjuntos de datos. Todos los conjuntos de datos deben tener las mismas categorías de modo que puedan mostrarse en la misma gráfica. La figura 3.18 es una gráfica de barras múltiple que muestra la cantidad de dinero que los hijos reciben de sus padres. Las categorías son los rangos de edad de los hijos. Los tres conjuntos de barras representan ingreso semanal de tres fuentes paternas diferentes: asignación (mesada), donativo y pago por tareas domésticas. (Nota: La mediana, que estudiaremos en la sección 4.1, es la mitad de un conjunto de valores; por ejemplo, la asignación mediana de \$3.00 para los hijos en edad de 9-10 significa que la mitad de estos niños recibieron más de \$3.00 y la mitad recibió menos de \$3.00).

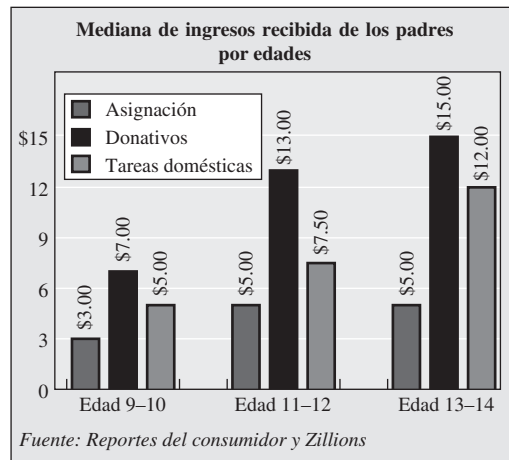


Figura 3.18 Una gráfica de barras múltiple. Fuente: Wall Street Journal Almanac.

Una **gráfica de líneas múltiple** sigue las mismas ideas básicas que la gráfica de barras múltiple, pero muestra los conjuntos de datos relacionados con líneas en lugar de barras. La figura 3.19 muestra una gráfica de líneas múltiple de precios de acciones, de bonos y de oro durante un periodo de 12 semanas. Esta gráfica particular también es una gráfica de serie de tiempo, ya que muestra datos sobre el tiempo.

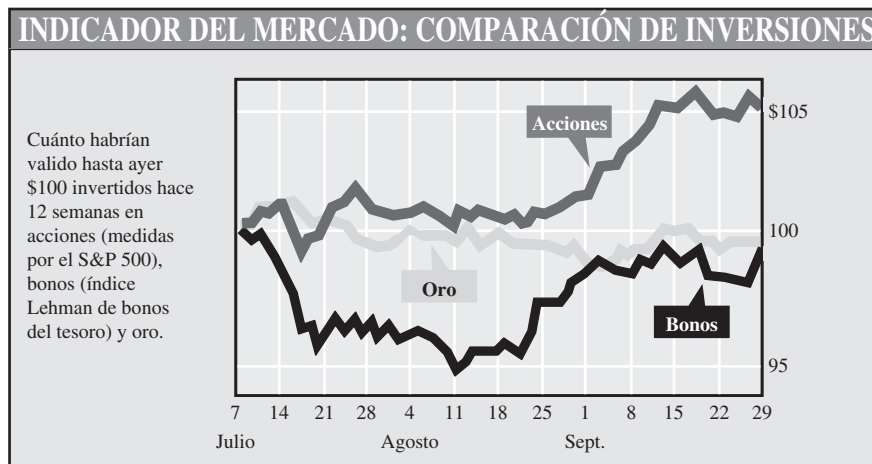


Figura 3.19 Una gráfica de líneas múltiple. Fuente: New York Times.

A propósito...

Alguna vez, el oro fue considerado una inversión sólida y parte importante de cualquier portafolio de inversión. Sin embargo, los precios del oro languidecieron en décadas recientes. Al final de 2006 el oro tenía un valor de sólo \$630 por onza, mucho menos que su valor ajustado por la inflación de más de \$2000 por onza en 1980.

EJEMPLO 1 Lectura de la gráfica de inversión

Considere la figura 3.19. Suponga que el 7 de julio había invertido \$100 en un fondo de acciones que sigue a S&P 500, \$100 en un fondo de bonos que sigue al Índice Lehman y \$100 en oro. Si el 15 de septiembre vende los tres fondos, ¿cuánto habría ganado o perdido?

Solución La gráfica muestra que los \$100 en el fondo de acciones habrían tenido un valor de alrededor de \$105 el 15 de septiembre. La inversión de \$100 en bonos habría disminuido a alrededor de \$99. La inversión en oro habría conservado su valor inicial de \$100. El 15 de septiembre su portafolio completo habría valido

$$\$105 + \$99 + \$100 = \$304$$

Usted habría ganado \$4 sobre su inversión total de \$300.

Apropósito...

En 1993 sólo 3 millones de personas en todo el mundo estaban conectadas a internet. En 2007 el acceso a internet fue compartido por más de 210 millones de estadounidenses y más de mil millones de personas en todo el mundo.

EJEMPLO 2 Conversión de gráficas

La figura 3.20 es una gráfica de barras múltiple de los números de hogares en Estados Unidos con computadoras y el número de hogares en línea (o conectados a internet). Vuelva a dibujar la gráfica como una gráfica de líneas múltiple. Analice brevemente las tendencias que muestran las gráficas.

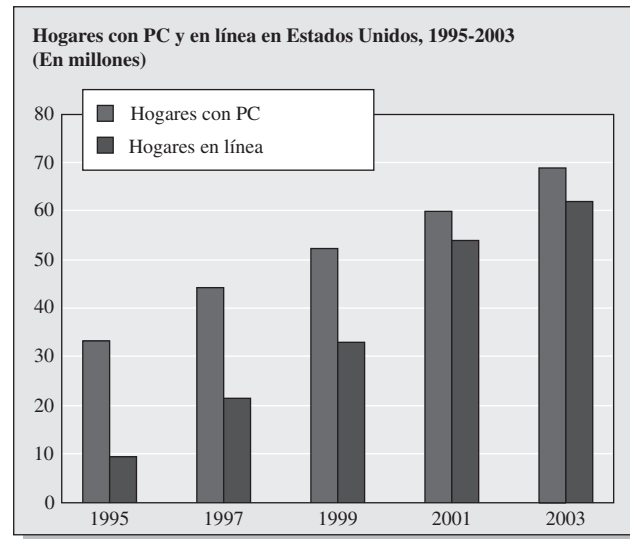


Figura 3.20 Una gráfica de barras múltiple de las tendencias de cómputo en casa. Fuente: *Resumen Estadístico de Estados Unidos*.

Solución Para convertir la gráfica a una de líneas múltiple, debemos cambiar los dos conjuntos de datos de barras a un conjunto de dos líneas. Colocamos un punto que corresponda a la altura de cada barra en el *centro* de cada categoría (año) y luego unimos los puntos del mismo color con una línea de ese color, como se muestra en la figura 3.21.

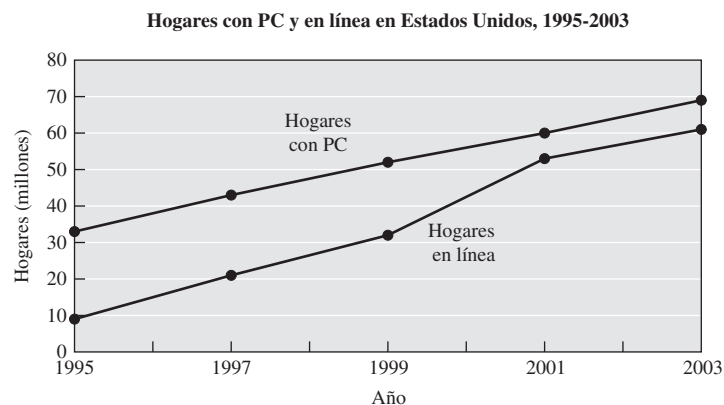


Figura 3.21 Diagrama de líneas múltiple que muestra los datos de la figura 3.20.

La tendencia más obvia es que ambos conjuntos de datos muestran un aumento con el tiempo. Vemos una segunda tendencia comparando las barras en cada año. En 1995 el número de hogares en línea (alrededor de 10 millones) era menor a un tercio del número de hogares con computadora (alrededor de 33 millones). En 2003 el número de hogares en línea (alrededor de

62 millones) fue casi 90% del número de hogares con computadora (alrededor de 70 millones). Esto nos dice que un porcentaje más alto de usuarios de computadora están poniéndose en línea. Si proyectamos las tendencias al futuro, parece probable que el número de hogares en línea se aproximará al número de hogares con computadora, y ambos se aproximarán al número total de hogares en Estados Unidos.

Gráficas apiladas

Otra forma de mostrar de manera simultánea dos o más conjuntos de datos relacionados es con una **gráfica apilada**, que muestra conjuntos de datos diferentes en una pila vertical. Aunque los datos puedan estar apilados en las gráficas de barras y en las gráficas de líneas, estas últimas son mucho más comunes.

EJEMPLO 3 Gráfica apilada de líneas

La figura 3.22 muestra las tasas de mortalidad (muertes por 100 000 personas) para cuatro enfermedades desde 1900. Con base en esta gráfica, ¿cuál fue la tasa de mortalidad por enfermedad cardiovascular en 1980? Analice las tendencias generales visibles en esta gráfica.

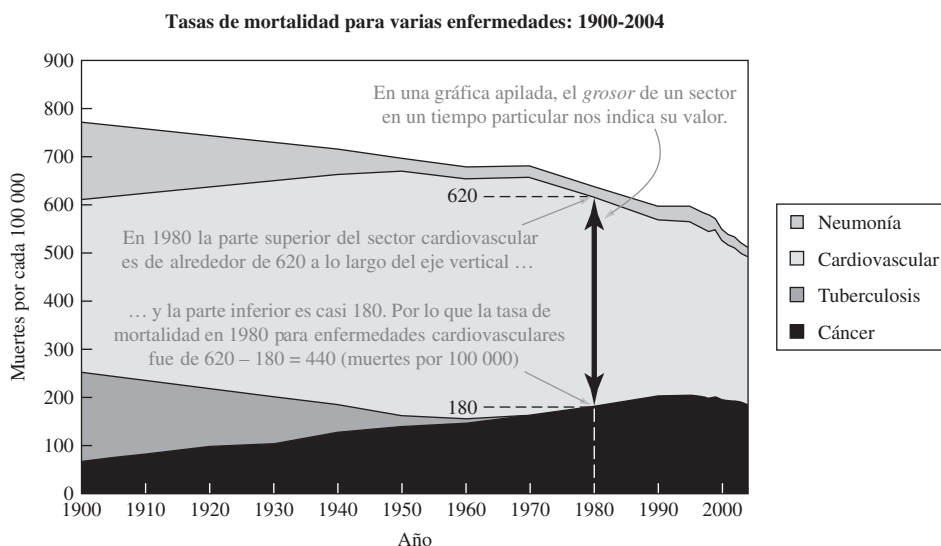


Figura 3.22 Una gráfica apilada usando sectores apilados. Fuentes: Centro Nacional para Estadísticas de Salud y Sociedad Americana del Cáncer.

A propósito...

La tasa de mortalidad del cáncer tuvo un aumento consistente en casi todo el siglo xx, alcanzando un pico en 1995 de alrededor de 205 muertes por 100 000 personas. Sin embargo, a partir de ese momento, la tasa de mortalidad del cáncer ha disminuido casi 10%, y en 2006 el número real de muertes por cáncer (no sólo la tasa por 100 000) disminuyó por primera vez. Los científicos atribuyen la disminución principalmente a una disminución en el número de personas que fuman (se estima que fumar provoca alrededor de 1/3 de todos los tipos de cáncer) y a los adelantos en la detección temprana y tratamiento del cáncer.

Solución Cada enfermedad tiene su propia región, o sector, coloreada; observe la importancia de la leyenda. El grosor de un sector en un tiempo particular nos indica su valor en ese tiempo. Para 1980, el sector cardiovascular se extiende de alrededor de 180 a 620 en el eje vertical, de modo que su grosor es alrededor de 440. Esto nos indica que la tasa de mortalidad en 1980 para enfermedades cardiovasculares fue de casi 440 muertes por 100 000 personas. La gráfica muestra varias tendencias importantes. Primera, la pendiente descendente del sector superior muestra que la tasa global de mortalidad de estas cuatro enfermedades disminuyó sustancialmente, de casi 800 muertes por 100 000 en 1900 a alrededor de 525 en 2003. La disminución radical en el grosor del sector de la tuberculosis muestra que esta enfermedad fue una gran causa de mortandad, pero casi ha sido erradicada desde 1950. Mientras tanto, el sector del cáncer muestra que la tasa de mortalidad del cáncer se elevó de manera constante hasta mediados de la década de los noventa, pero ha caído desde entonces.

Datos geográficos

Los datos del uso de energía en la tabla 3.3 son un ejemplo de **datos geográficos**, ya que los datos corresponden a diferentes lugares geográficos. Utilizamos estos datos para hacer una tabla de frecuencias (tabla 3.4), un histograma (figura 3.8) y un diagrama de tallos y hojas (figura 3.9). Sin embargo, éstos no nos dan un sentido de algún patrón geográfico en los datos. Por ejemplo, no somos capaces de ver si los estados al noreste tienen niveles similares de uso de energía. La figura 3.23 muestra una manera de presentar las tendencias geográficas. Las categorías se codifican con colores de acuerdo con la clave y cada estado se colorea de manera apropiada.

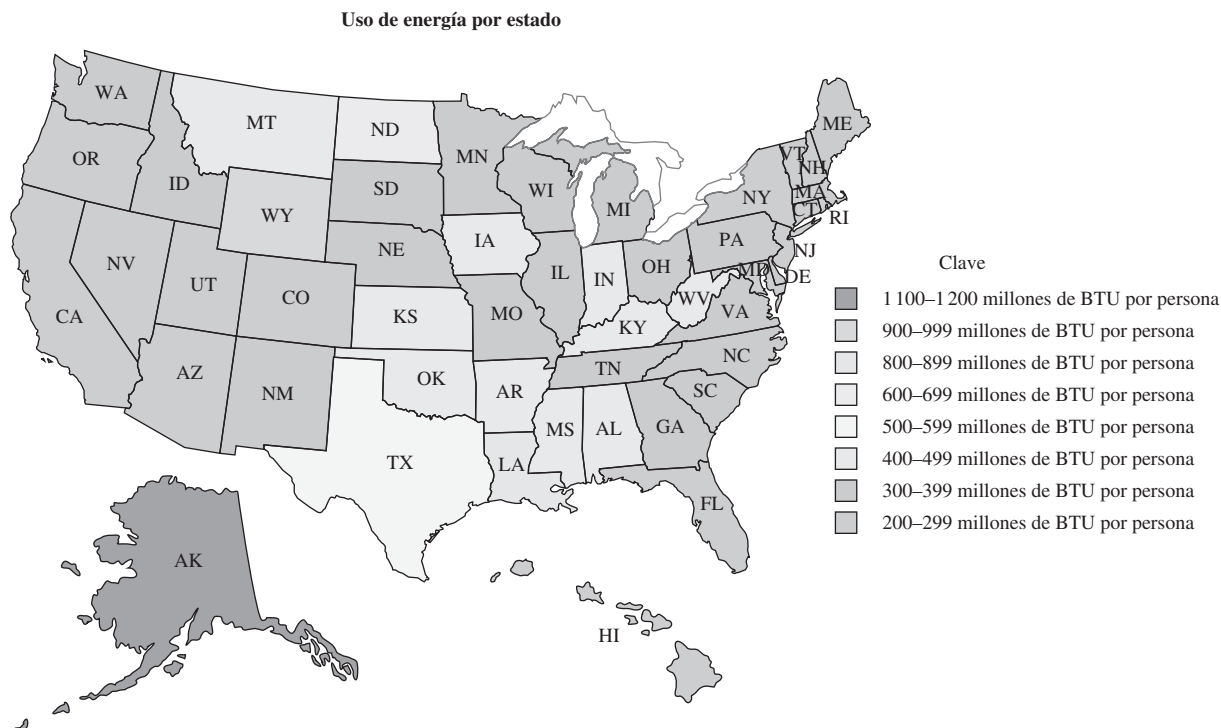


Figura 3.23 Datos geográficos pueden presentarse con un mapa codificado por colores.

UN MOMENTO DE REFLEXIÓN

¿Qué puede aprender del histograma de la figura 3.8 que no pueda aprender con facilidad de la presentación geográfica en la figura 3.23? ¿Qué puede aprender de la presentación geográfica que no pueda hacerlo del histograma? ¿Observa algunas tendencias geográficas sorprendentes en la figura 3.23? Explique.

La figura 3.23 funciona bien para el conjunto de datos de energía, ya que a cada estado corresponde un número único. Para datos que varían de manera continua a lo largo de áreas geográficas, es más conveniente un **mapa de contornos**. La figura 3.24 muestra un mapa de contornos de la temperatura en Estados Unidos en un momento particular. Cada uno de los *contornos* (líneas) conecta localidades con la misma temperatura. Por ejemplo, la temperatura es 50°F en todos lados a lo largo del contorno etiquetado con 50° y 60°F donde se etiqueta la línea con 60°. Entre estos dos contornos la temperatura está entre 50°F y 60°F. Observe que diferencias de temperaturas mayores significan contornos más espaciados. Por ejemplo, los contornos muy juntos en el noreste indican que los valores de la temperatura varían de manera sustancial a pequeñas distancias. Para hacer la gráfica más sencilla de leer, las regiones entre contornos adyacentes se codifican con un color.

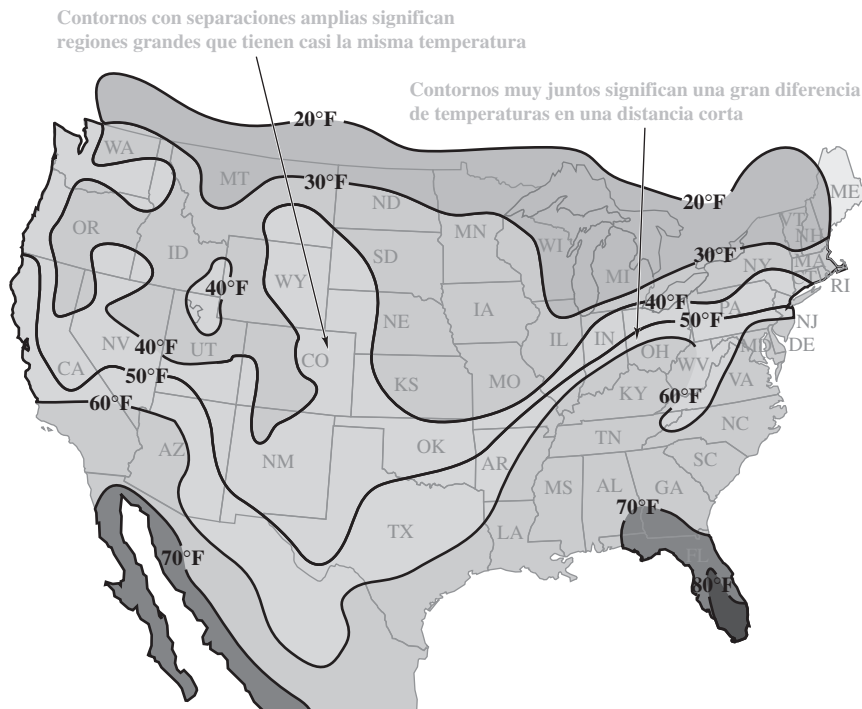


Figura 3.24 Datos geográficos que varían de manera continua, tal como las temperaturas, pueden mostrarse con un mapa de contornos.

EJEMPLO 4 Mapa de contornos de nivel

Los diagramas de contornos también se utilizan para mostrar altitudes (o niveles) geográficas. La figura 3.25 muestra contornos de altitudes alrededor de Boulder, Colorado. Analice unas cuantas características clave que muestra el mapa.

Solución Las etiquetas en la figura indican las características clave. Los contornos están ampliamente espaciados en el este, donde el terreno es relativamente plano y las altitudes son casi constantes. Hacia el oeste los contornos son muy densos, donde las montañas se elevan desde las planicies. Los contornos concéntricos en el centro del mapa rodean cimas de montañas.

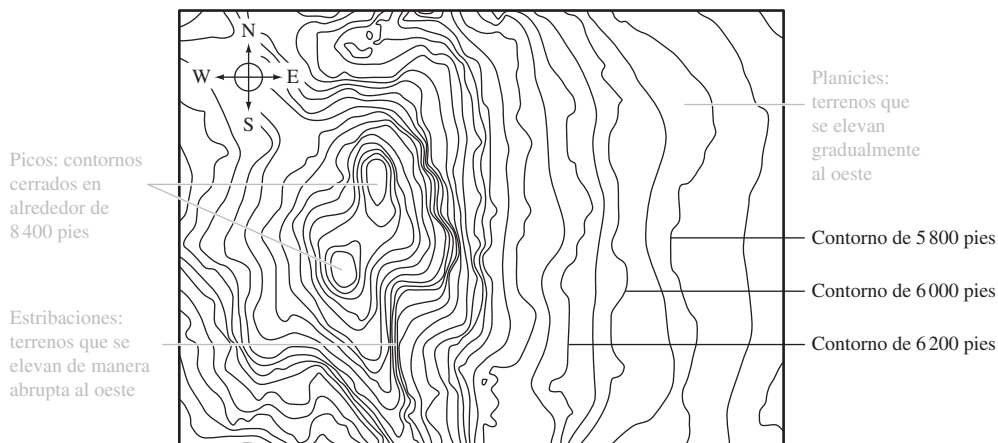


Figura 3.25 Un mapa de contornos de altitud para la región cercana a Boulder, Colorado.

Gráficas en tres dimensiones

En la actualidad, los programas de cómputo hacen sencillo dar a casi cualquier gráfica una apariencia de tres dimensiones. Por ejemplo, la figura 3.26 muestra la misma gráfica de barras que la figura 3.1, pero “arreglada” con una apariencia de tres dimensiones. Podría parecer bonita, pero el efecto de tres dimensiones es sólo decorativo; no proporciona información que no haya mostrado ya la figura 3.1.

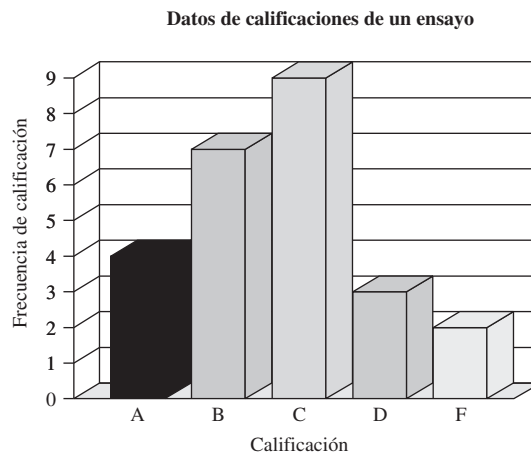


Figura 3.26 Esta gráfica tiene una apariencia de tres dimensiones, pero sólo muestra datos de dos dimensiones.

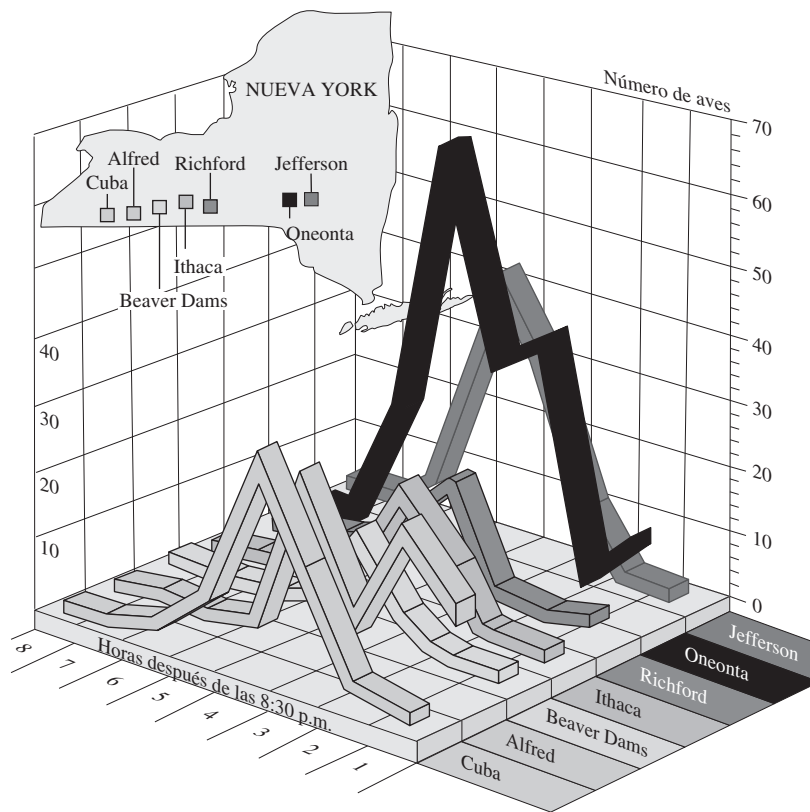
En contraste, cada uno de los tres ejes en la figura 3.27 lleva información distinta, da valor a la gráfica de tres dimensiones. Los investigadores estudian los patrones de migración de una especie de ave (el *charlatán*) contando el número de aves que pasan volando sobre siete ciudades de Nueva York por la noche. Como se muestra en el mapa inserto, las ciudades fueron alineadas de este a oeste de modo que los investigadores supieron sobre qué partes del estado vuelan las aves y en qué momentos de la noche, cuando ellas emigran hacia el sur en invierno. Observe que los tres ejes miden *número de aves*, *hora en la noche* y *ubicación este-oeste*.

EJEMPLO 5 Migración de aves

Con base en la figura 3.27, ¿aproximadamente a qué hora voló la mayor cantidad de aves sobre la línea este-oeste marcada por las siete ciudades? ¿En qué parte de Nueva York voló la mayoría de las aves? Aproximadamente, ¿cuántas aves pasaron sobre Oneonta alrededor de las 12:30 a.m.?

Solución El número de aves detectado en todas las ciudades tuvo un pico entre 4 y 6 horas, después de las 8:30 p.m., o entre las 12:30 y las 2:30 a.m. Una mayor cantidad de aves voló sobre las dos ciudades de más al este, Oneonta y Jefferson, que sobre las ciudades de más al oeste, lo cual significa que la mayoría de las aves pasaron volando sobre la parte este del estado. Para responder la pregunta específica acerca de Oneonta, observe que 12:30 a.m. son 4 horas después de las 8:30 p.m. En la gráfica esta hora se alinea con la hondonada entre los picos en la línea de Oneonta. Buscando en el eje *número de aves*, vemos que entre 25 y 30 aves pasaron volando sobre Oneonta a esa hora.

SEGUIMIENTO SÓNICO DE RUTAS DE MIGRACIÓN DE AVES



Fuente: Bill Evans/
Laboratorio de Ornitología de Cornell

Figura 3.27 Esta gráfica muestra verdaderos datos de tres dimensiones. Fuente: *New York Times*.

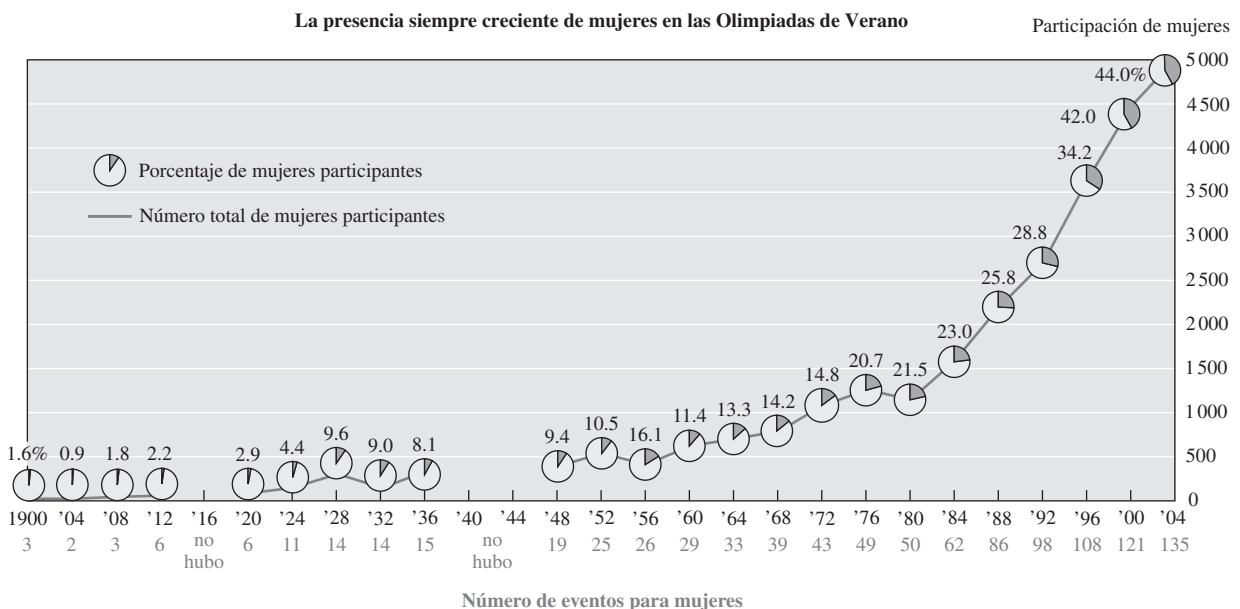
Gráficas combinadas

Todos los tipos de gráficas que hemos estudiado hasta aquí son comunes y muy fáciles de crear. Pero ahora los medios están llenos con muchas variedades de gráficas más complejas. Por ejemplo, la figura 3.28 muestra una gráfica concerniente a la participación de mujeres en los juegos olímpicos de verano. Esta sola gráfica combina una gráfica de líneas, muchas gráficas circulares y datos numéricos. En efecto, es un caso de “un dibujo dice más que mil palabras”.

EJEMPLO 6 Mujeres olímpicas

Describe tres tendencias que se muestran en la figura 3.28, en la página 120.

Solución La gráfica de líneas muestra que el número total de mujeres que compiten en las olimpiadas de verano ha crecido de manera consistente, en especial desde 1960, llegando a 5000 en los juegos de 2004. Las gráficas circulares muestran que el porcentaje de mujeres entre todos los competidores también ha aumentado, alcanzando 44% en los juegos de 2004. Los números en la parte inferior muestran que el número de eventos en que las mujeres compiten también se ha incrementado radicalmente, llegando a 135 en los juegos de 2004.



Fuente: Comité Olímpico Internacional.

Figura 3.28 Mujeres en las Olimpiadas. Adaptado del *New York Times*.

UN MOMENTO DE REFLEXIÓN

¿Cuál de las tendencias mostradas en la figura 3.28 es probable que continúe durante los siguientes juegos olímpicos? ¿Cuáles no? Explique.

Sección 3.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Gráficas en tres dimensiones.** Una revista utilizó barriles de petróleo en tres dimensiones de distintos tamaños para presentar la producción del combustible de diferentes países. Los barriles de petróleo, claramente, eran objetos de tres dimensiones pero ¿las cantidades de producción de petróleo de diferentes países son ejemplos de datos de tres dimensiones? ¿Los barriles de petróleo son apropiados para estos datos? Si no, ¿cómo deben representarse los datos?
- Gráficas apiladas.** ¿Qué son las gráficas apiladas y cómo ayudan?
- Datos geográficos.** ¿Qué son datos geográficos? Identifique al menos dos formas de presentar datos geográficos.
- Mapa de contornos.** ¿Qué es un contorno en un mapa de contorno? ¿Qué quiere decir que los contornos estén muy juntos unos de otros? ¿Qué significa que estén muy separados?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Gráficas en tres dimensiones.** Sue asegura que necesita una gráfica en tres dimensiones para mostrar la temperatura máxima diaria en Sacramento del año pasado.
- Gráfica apilada de líneas.** Una gráfica apilada de líneas podría usarse para presentar los números de hombres y mujeres en su universidad en los 10 años anteriores.
- Mapa de contornos.** Un mapa de contornos podría usarse para presentar la población en cada ciudad importante de Estados Unidos.
- Datos geográficos.** Un artista gráfico para una revista está mostrando la población de las ciudades principales de Estados Unidos usando barras de alturas diferentes, con las barras colocadas en las ubicaciones de las ciudades en un mapa de Estados Unidos.

Conceptos y aplicaciones

- Género de estudiantes.** La gráfica apilada en la figura 3.29 muestra los números de estudiantes hombres y mujeres en educación superior para diferentes años. Las proyecciones son del Centro Nacional para Estadísticas de Educación de Estados Unidos.
 - En palabras, analice las tendencias reveladas en esta gráfica.

- b. Haga nuevamente la gráfica como una de líneas múltiple. Analice brevemente las ventajas y desventajas de las dos representaciones de este conjunto particular de datos.

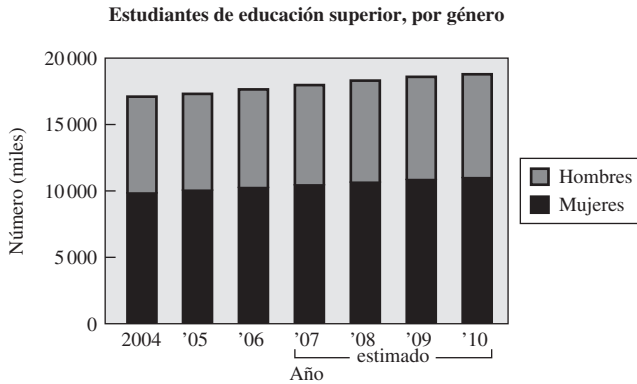


Figura 3.29

10. Precios de casas por región. La gráfica en la figura 3.30 muestra los precios de las casas en regiones diferentes de Estados Unidos. Observe que los datos *no* se han ajustado por los efectos de la inflación.

- Describa las tendencias generales que se aplican a los datos de precios de casas para todas las regiones.
- Describa las diferencias que observa entre las diversas regiones.
- Describa cómo esta gráfica se vería diferente si los datos fuesen ajustados por los efectos de la inflación.

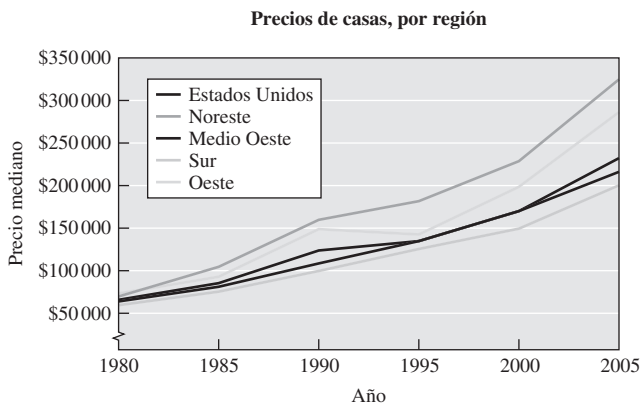


Figura 3.30

11. Género e ingreso. Considere la gráfica en la figura 3.31 de los ingresos medianos de hombres y mujeres en diferentes años (Oficina del Censo de Estados Unidos).

- ¿Qué dice la gráfica? ¿La razón de ingresos de hombres en comparación con los ingresos de mujeres parece que está decreciendo? ¿Por qué sí o por qué no?

- b. Vuelva a dibujar la gráfica como una gráfica de líneas múltiple (dos). Brevemente analice las ventajas y desventajas de las dos representaciones de este conjunto particular de datos.

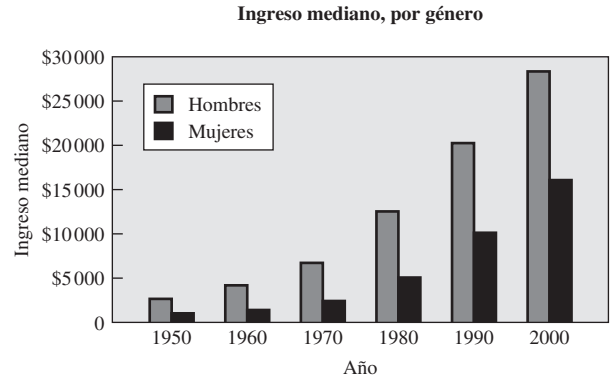


Figura 3.31

12. Tasas de matrimonios y divorcios. La gráfica en la figura 3.32 muestra las tasas de matrimonios y divorcios en Estados Unidos para ciertos años desde 1900. Ambas tasas se dan en unidades de matrimonios o divorcios por 1000 personas en la población (Departamento de Salud y Servicios Humanos).

- ¿Por qué estos datos consisten en tasas de matrimonios y divorcios en lugar de números totales de matrimonios y divorcios? Comente cualesquiera tendencias que observe en estas tasas y proporcione explicaciones históricas y sociológicas plausibles para estas tendencias.
- Construya una gráfica apilada de los datos de tasa de matrimonio y tasa de divorcio. Para cada barra coloque la tasa de divorcio arriba de la tasa de matrimonio. ¿Cuál gráfica hace más sencilla las comparaciones: la gráfica de barras múltiple mostrada aquí o la gráfica apilada? Explique.

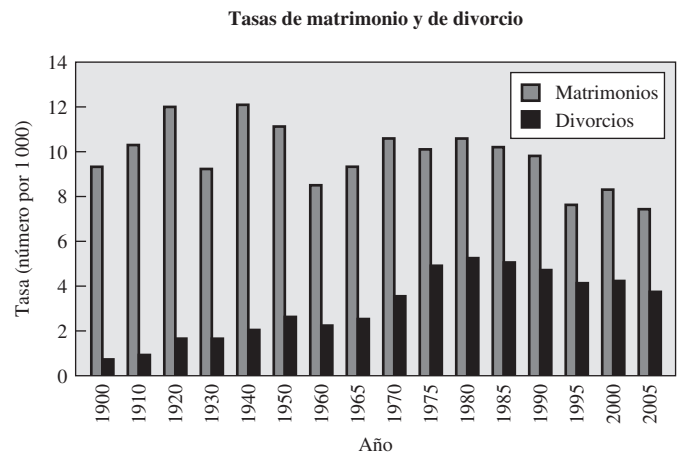


Figura 3.32

13. Gasto federal. La gráfica de líneas apiladas en la figura 3.33 muestra los cambios en las principales categorías de gastos del presupuesto federal. Observe que la categoría “interés neto” representa los pagos de interés en la deuda nacional, y la categoría “todo lo demás” incluye el gasto en cosas tales como educación, cuidado ambiental e investigación científica. Interprete esta gráfica y analice algunas de las cosas que revela.

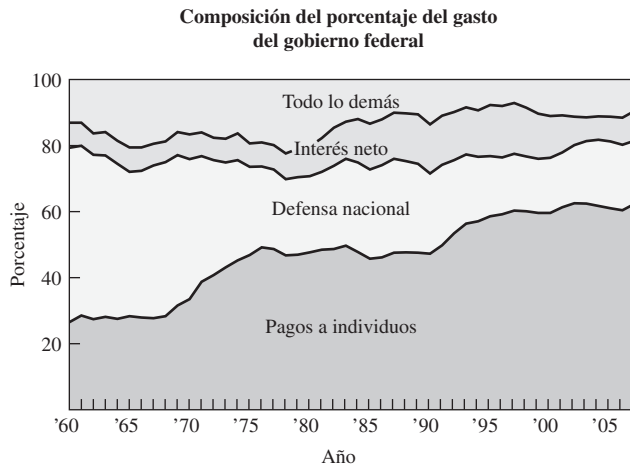


Figura 3.33 Fuente: Oficina de Administración y Presupuesto.

14. Grados universitarios. La gráfica de líneas apiladas en la figura 3.34 muestra los números de graduados universitarios para hombres y mujeres a lo largo del tiempo.

- a. Estime los números de graduados universitarios para hombres y mujeres (de manera separada) en 1930 y 2000.

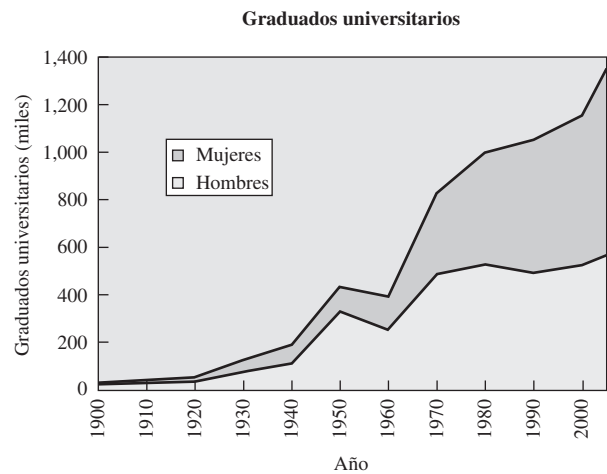


Figura 3.34

- b. Compare los números de graduados para hombres y mujeres (por separado) en 1980 y 2000.
- c. ¿Durante qué década el número *total* de graduados aumentó más?
- d. Compare los números *totales* de graduados en 1950 y 2000.
- e. ¿Considera que la gráfica de líneas apiladas es una manera efectiva de mostrar estos datos? Analice brevemente otras formas que podrían haberse usado en su lugar.

15. Mortandad debida al melanoma. La figura 3.35 muestra cómo la tasa de mortandad debida al *melanoma* (una forma de cáncer en la piel) varía de condado a condado en todo Estados Unidos. La leyenda muestra que entre más oscuro

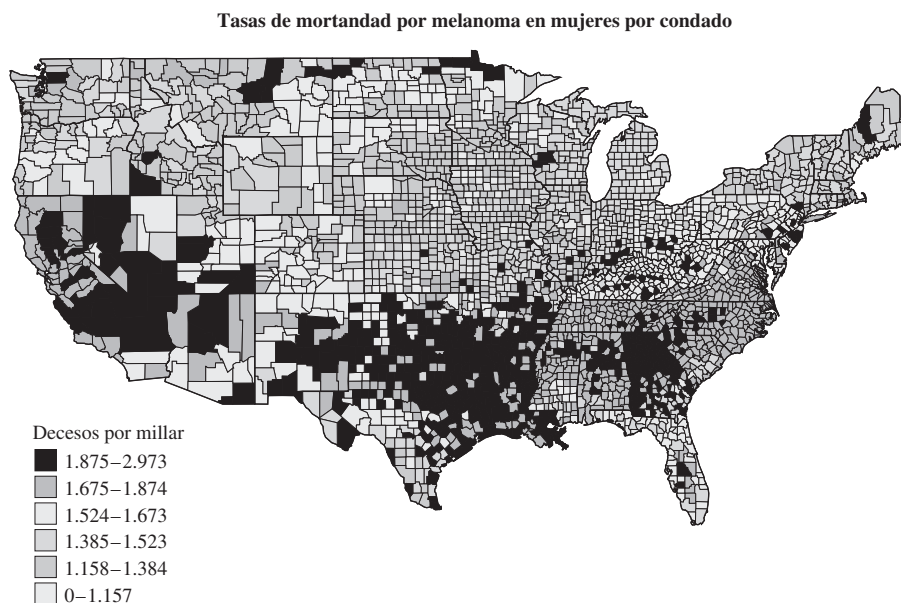
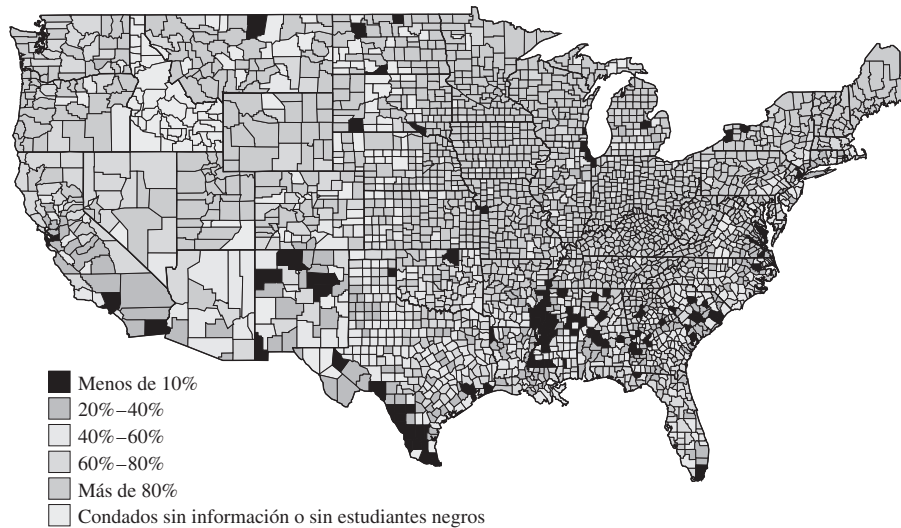


Figura 3.35 Fuente: Profesor Karen Kafadar, Departamento de Matemáticas, Universidad de Colorado en Denver.

Probabilidad de que un estudiante negro tenga compañeros de clase blancos

Figura 3.36 Fuente: *New York Times*.

sea el sombreado en un condado más alta es la tasa de mortandad. Analice algunas tendencias que revela la figura. Si usted estuviese investigando el cáncer de piel, ¿cuáles regiones merecerían estudio especial? ¿Por qué?

- 16. Segregación escolar.** Una manera de medir la segregación es la probabilidad de que un estudiante negro tendrá compañeros que son blancos. Un estudio reciente del *New York Times* encontró que, con esta medida, la segregación aumentó de manera significativa en la década de los noventa. La figura 3.36 muestra la probabilidad de que un estudiante negro tenga compañeros blancos, por condado, durante un año escolar reciente. ¿Parecen existir diferencias regionales significativas? ¿Puede reconocer alguna diferencia entre áreas urbanas y áreas rurales? Analice explicaciones posibles para algunas tendencias que vea en la figura.

Creación de figuras. Los ejercicios 17 a 20 proporcionan tablas de datos reales. Para cada tabla construya una presentación gráfica de los datos. Puede elegir cualquier tipo de gráfica que sienta que es apropiada para el conjunto de datos. Además, para hacer la presentación, escriba unas cuantas oraciones que expliquen por qué eligió este tipo de presentación y unas cuantas oraciones que describan patrones interesantes en los datos.

- 17. Alcohol en la carretera.** La tabla siguiente proporciona el número de accidentes automovilísticos mortales en los que (a) no estuvo involucrado el alcohol, (b) estuvo involucrado alcohol en pequeña cantidad (contenido de alcohol en la sangre entre 0.01 y 0.09) y (c) estuvo involucrado un alto nivel de alcohol (contenido de alcohol en la sangre por arriba de 0.1). Todas las cifras están en miles de muertes (Departamento Nacional de Seguridad en Carreteras).

Año	Sin alcohol	Poco alcohol	Mucho alcohol
1982	18.7	4.8	20.4
1984	20.5	4.8	19.0
1986	22.0	5.1	18.9
1988	23.5	4.9	18.7
1990	22.5	4.4	17.7
1992	21.4	3.6	14.2
1994	24.1	3.5	13.1
1996	24.8	3.8	13.4
1998	25.5	3.7	12.2
2000 (est.)	25.1	3.9	12.8

- 18. Periódicos.** La tabla siguiente proporciona el número de periódicos y su circulación total (en millones) para años seleccionados desde 1920 (*Editor & Publisher*)

Año	Número de periódicos	Circulación (millones)
1920	2042	27.8
1930	1942	39.6
1940	1878	41.1
1950	1772	53.9
1960	1763	58.8
1970	1748	62.1
1980	1747	62.2
1990	1611	62.3
2000	1485	56.1
2003	1456	55.2

19. Víctimas de arma de fuego. La tabla siguiente resume las muertes debidas a armas de fuego en países diferentes en un año reciente (Coalición para la detención de violencia con armas de fuego).

País	Total de víctimas por arma de fuego	Homicidios con arma de fuego	Suicidios por arma de fuego	Accidentes fatales por arma de fuego
Estados Unidos	35 563	15 835	18 503	1 225
Alemania	1 197	168	1 004	25
Canadá	1 189	176	975	38
Australia	536	96	420	20
España	396	76	219	101
Reino Unido	277	72	193	12
Suiza	200	27	169	4
Vietnam	131	85	16	30
Japón	93	34	49	10

20. Madres en la fuerza laboral. La tabla siguiente lista los porcentajes de madres en la fuerza laboral que tienen hijos en dos diferentes rangos de edades (con base en datos de la Oficina de Estadísticas de Trabajo).

Años	6 a 17 años	Menos de 6
1955	38.4	18.2
1965	45.7	25.3
1975	54.9	39.0
1985	69.9	53.5
1995	76.4	62.3
2005	77.5	62.2

Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 3 en www.aw.com/bbt.

- 21. Mapas de clima.** Muchos sitios web ofrecen mapas de contornos con datos de clima actuales. Por ejemplo, puede utilizar el sitio Yahoo! Weather para generar muchos mapas de contorno de climas. Genere al menos dos mapas de contornos de climas y analice los que muestran.
- 22. Presupuesto federal.** Vaya al sitio web de la Oficina de Administración y Presupuesto de Estados Unidos y busque algunas de sus gráficas relacionadas al presupuesto federal. Elija dos gráficas que le interesen y analice qué muestran los datos.
- 23.** Consulte la figura 3.37 en la que los porcentajes están representados por volúmenes de partes de la cabeza de alguien. ¿Los datos están presentados en un formato que los hace fáciles de entender y comparar? ¿Los datos están presentados de manera que no causen confusión? ¿La misma

información puede presentarse de una menor manera? Si es así, construya su propia gráfica que presente de mejor manera la información dada.

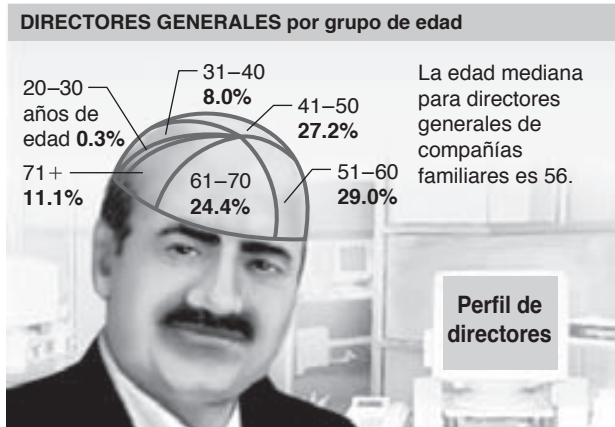


Figura 3.37 Fuente: Datos de Arthur Anderson/Encuesta de 1997 Mass Mutual Family Business.

EN LAS NOTICIAS

- 24. Gráficas de barras múltiples.** Determine un ejemplo de una gráfica de barras múltiple o gráfica de líneas múltiple en un reporte reciente. Comente sobre la efectividad de la presentación. ¿Otra presentación podría haber sido usada para mostrar los mismos datos?
- 25. Gráficas apiladas.** Encuentre un ejemplo de una gráfica apilada en una noticia reciente. Comente sobre la efectividad de la presentación. ¿Otra presentación podría haber sido usada para mostrar los mismos datos?
- 26. Datos geográficos.** Encuentre un ejemplo de una gráfica de datos geográficos en una noticia reciente. Comente sobre la efectividad de la presentación. ¿Otra presentación podría haber sido usada para mostrar los mismos datos?
- 27. Presentaciones en tres dimensiones.** Encuentre un ejemplo de una presentación en tres dimensiones en una noticia reciente. ¿Las tres dimensiones son necesarias, o están incluidas por razones decorativas? Comente sobre la efectividad de la presentación. ¿Otra presentación podría haber sido usada para mostrar los mismos datos?
- 28. Gráficas sofisticadas en noticias.** Encuentre un ejemplo en las noticias de una gráfica que combine dos o más de los tipos de gráficas básicas. Explique brevemente qué muestra la gráfica y analice la efectividad de la gráfica.

3.4 Algunas precauciones con respecto a gráficas

Como hemos visto, las gráficas pueden ofrecer resúmenes claros y significativos de datos estadísticos. Sin embargo, incluso gráficas muy bien realizadas pueden ser engañosas, si no se tiene cuidado al interpretarlas, y gráficas pobremente hechas casi siempre son confusas. Además, algunas personas utilizan las gráficas de formas deliberadamente engañosas. En esta sección analizamos unas de las maneras más comunes en las cuales las gráficas pueden llevar a confundirnos.

Distorsiones sensoriales

Muchas gráficas se dibujan de manera que distorsionan nuestra percepción de ellas. La figura 3.38 muestra uno de los tipos más comunes de distorsión. Las barras en forma de dólar se utilizan para mostrar la disminución del valor del dólar con el tiempo. Las *longitudes* de las barras representan los datos, pero nuestros ojos tienden a centrarse en las *áreas* de las barras. Por ejemplo, la barra de la derecha tiene la intención de mostrar que un dólar en 2006 tenía un valor de 41% del valor de un dólar en 1980. Su longitud en realidad es 41% de la de la barra de la izquierda, pero su área es mucho más pequeña en comparación (alrededor de 17% del área de la barra de la izquierda). Esto da la percepción de que el valor del dólar se contrajo mucho más de lo que en realidad lo hizo.

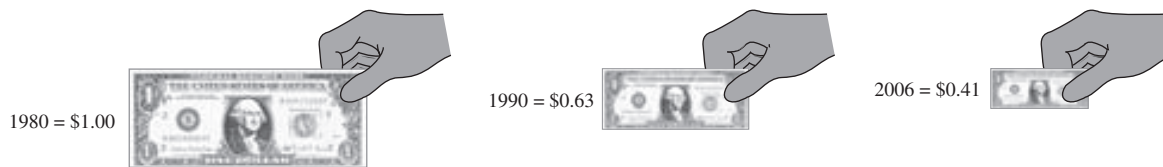


Figura 3.38 Las longitudes de los dólares son proporcionales a su poder de compra, pero nuestros ojos se desvían a las áreas, las cuales disminuyen más que las longitudes.

Distorsiones aún mayores ocurren cuando una gráfica muestra volumen donde la medida importante es la longitud. La figura 3.39 utiliza televisores para representar el número de hogares con televisión por cable en 1980 y 2005. Este número ha aumentado en un factor de alrededor de 4 durante ese periodo, de 18 millones a 73 millones de hogares, como se muestra por medio de las *alturas* de los televisores en la figura. Sin embargo, nuestros ojos son llevados a los *volúmenes* de los televisores, que difieren en un factor mucho mayor de alrededor de $4^3 = 64$ y, por tanto, hace aumentar la apariencia mucho más de lo que en realidad fue.

Hogares con TV por cable



Figura 3.39 Las alturas de los televisores son la medida importante en esta figura, pero nuestros ojos se desvían a sus volúmenes.

A propósito...

Investigadores alemanes a finales del siglo XIX estudiaron muchos tipos de gráficas. El tipo de distorsión en las figuras 3.38 y 3.39 fue tan común que le dieron su propio nombre, que se traduce más o menos como “el viejo truco de resaltar el efecto reduciendo la visión”.

A propósito...

El crecimiento de la televisión por cable ha disminuido en los años recientes a consecuencia de la competencia de los servicios de televisión satelital. La competencia puede aumentar aún más cuando las líneas telefónicas e internet sean capaces de proporcionar programación de televisión.

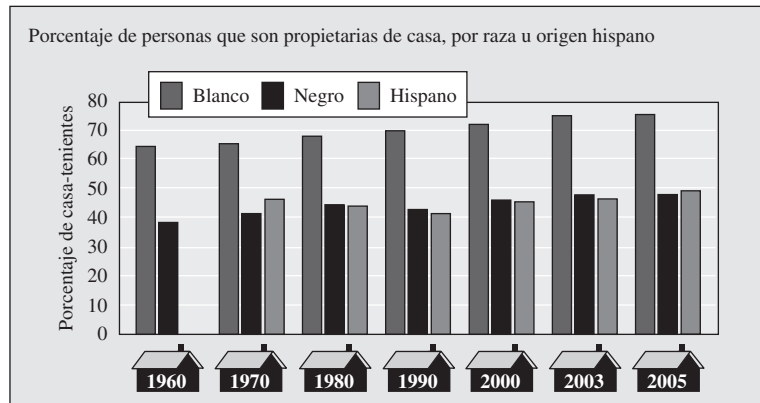


Apropósito...

En 2005 el porcentaje general de hogares en Estados Unidos ocupado por sus propietarios había alcanzado 69.0%, el porcentaje más alto en la historia.



Hogar, dulce hogar



Fuente: Oficina del Censo de Estados Unidos

Figura 3.40 Casa-tenientes por raza u origen hispano. Observe la escala horizontal no uniforme. Fuente: Departamento de Comercio de Estados Unidos, Oficina del Censo de Estados Unidos; adaptado del *Wall Street Journal Almanac*.

Tenga cuidado con las escalas

La figura 3.40 muestra una gráfica de múltiples barras respecto a casa-tenencia en Estados Unidos. A primera vista parece que los porcentajes de personas propietarias de sus casas se elevaron mucho más rápido durante el periodo 1960-2000 que en años recientes. Pero observe la escala horizontal con mayor cuidado: las primeras cinco categorías representan años que están una década separados, mientras que las últimas dos categorías representan los años 2003 y 2005. Si los pequeños aumentos de 2003 a 2005 se repitiesen cada año para una década entera, el aumento no se vería muy diferente de las tendencias de las otras décadas. Esta gráfica es engañosa a primera vista, ya que no utiliza una escala uniforme para el eje horizontal.

UN MOMENTO DE REFLEXIÓN

Con base en la figura 3.40 y su intuición, pronostique las tasas de casa-tenencia en Estados Unidos (por raza u origen hispano) en 2020. Explique el razonamiento que respalda sus pronósticos.

Problemas semejantes de escalas pueden ocurrir con el eje vertical. La figura 3.41a muestra el porcentaje de estudiantes universitarios femeninos entre 1910 y 2005. A primera vista parece que este porcentaje creció muchísimo después de 1950. Pero la escala en el eje vertical no inicia en cero y no termina en 100%. El aumento sigue siendo sustancial pero parece lejos de ser drástico si volvemos a dibujar la gráfica con el eje vertical cubriendo el rango completo de 0 a 100% (figura 3.41b). Desde un punto de vista matemático, dejar fuera el punto cero en una escala es perfectamente honesto y puede hacer más sencillo ver tendencia en los datos a pequeña escala. Sin embargo, como muestra este ejemplo, puede ser engañoso si no estudiamos la escala con mucho cuidado.

En ocasiones la escala podría no ser engañosa, pero podría malinterpretarse a menos que la examine con cuidado. Considere la figura 3.42a, que muestra cómo las velocidades de las computadoras más rápidas se ha incrementado con el tiempo. A primera vista parece que las velocidades han aumentado en forma lineal. Por ejemplo, parecería como si la velocidad aumentara la misma cantidad de 1990 a 2000 que de 1950 a 1960. Sin embargo, si revisamos con cuidado, vemos que cada marca en la escala vertical representa un aumento de *diez veces* en velocidad. Ahora vemos que la velocidad de las computadoras creció alrededor de 1 a 100 cálculos por segundo de 1950 a 1960, y de alrededor de 100 millones a 10 mil millones de cálculos por segundo entre 1990 y 2000. Este tipo de escala se denomina *escala exponencial*.

La persona más fácil de engañar es uno mismo.

—Edward Bulwer-Lytton

Mujeres como un porcentaje de todos los estudiantes universitarios

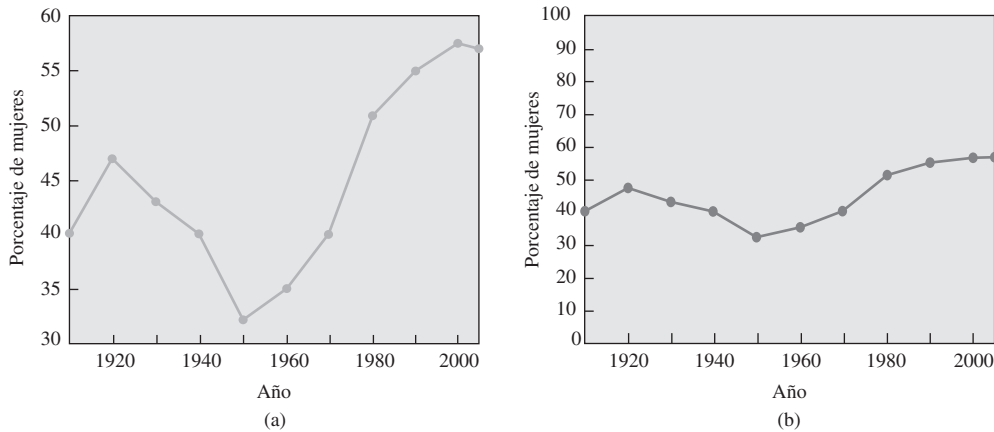


Figura 3.41 Ambas gráficas muestran los mismos datos, pero se ven muy diferentes porque sus escalas verticales tienen rangos diferentes.

(o *escala logarítmica*) ya que crece en potencias de 10 y las potencias de 10 tienen *exponentes*. (Por ejemplo, 3 es el exponente de 10^3). Las escalas exponenciales son útiles para presentar datos que varían en un enorme rango de valores, como puede ver al comparar la gráfica exponencial de la figura 3.42a con la gráfica común en la figura 3.42b. Puesto que las velocidades han crecido tan rápidamente, la escala común hace imposible ver cualquier detalle de los primeros años en la gráfica.

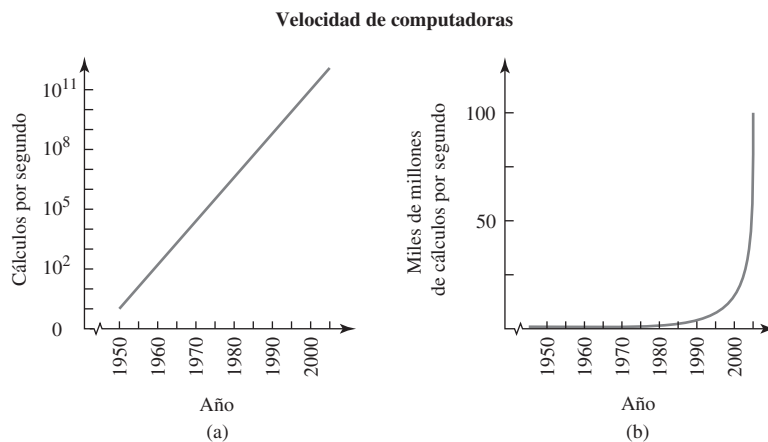


Figura 3.42 Ambas gráficas muestran los mismos datos, pero la gráfica de la izquierda utiliza una escala exponencial (logarítmica).

A propósito...

En 1965 el fundador de Intel, Gordon E. Moore, pronosticó que los avances en la tecnología permitirían a los chips de las computadoras duplicar la potencia aproximadamente cada dos años. Esta idea ahora se llama *ley de Moore*, y casi se ha cumplido desde que Moore la estableció por primera vez.

UN MOMENTO DE REFLEXIÓN

Con base en la figura 3.42a, ¿puede pronosticar la velocidad de las computadoras más rápidas en 2015? ¿Podría hacer el mismo pronóstico con la figura 3.42b? Explique.

ESTUDIO DE CASO

Amenaza de un asteroide

Los asteroides y los cometas en ocasiones chocan con la Tierra. Los pequeños tienden a consumirse en la atmósfera o a crear pequeños cráteres con el impacto. Pero los grandes podrían causar una devastación significativa. Hace alrededor de 65 millones de años, un asteroide de casi

Apropósito...

Más de 150 cráteres de impacto han sido identificados alrededor de la Tierra. Uno de los mejor conservados es el Cráter del Meteorito en Arizona (fotografía de abajo). Se formó hace 50 000 años cuando un asteroide de casi 50 metros se estrelló con la Tierra y liberó energía equivalente a la de una bomba de hidrógeno de 20 megatones y dejó un cráter de un kilómetro de largo y 200 metros de profundidad.

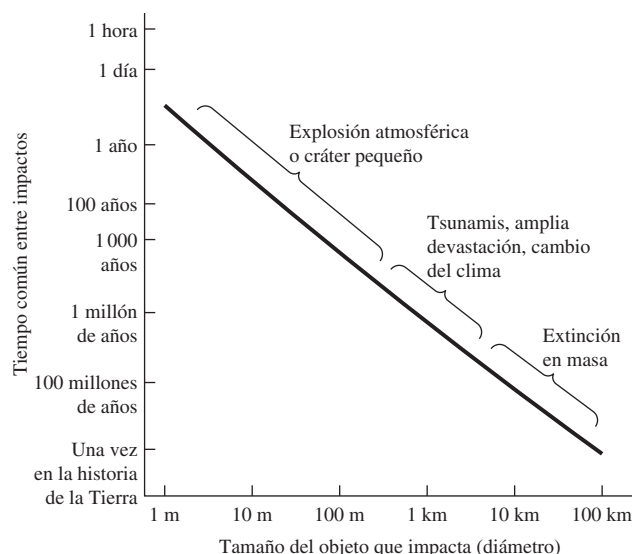


Figura 3.43 Esta gráfica muestra cómo la frecuencia de impactos —y la magnitud de sus efectos— dependen del tamaño del objeto que impacta. Observe que impactos menores son mucho más frecuentes que los grandes.

10 kilómetros de diámetro chocó con la Tierra, dejando un amplio cráter de 200 kilómetros en la costa de la península de Yucatán en México. Los científicos suponen que este impacto causó la extinción de alrededor de tres cuartos de todas las especies que vivían en la Tierra en ese tiempo, incluyendo a todos los dinosaurios.

Claro, un impacto similar serían malas noticias para nuestra civilización. Por tanto, podríamos querer entender la posibilidad de tal evento. La figura 3.43 muestra una gráfica que relaciona el tamaño de asteroides y cometas que han impactado con la frecuencia con que tales objetos chocan con la Tierra. A consecuencia del amplio rango de tamaños y escalas de tiempo involucradas, *ambos* ejes en esta gráfica son exponenciales. El eje horizontal muestra la frecuencia del impacto; moviéndose hacia arriba en el eje vertical corresponde a eventos más frecuentes. Con esta gráfica doblemente exponencial, podemos ver con claridad las tendencias. Por ejemplo, objetos pequeños de alrededor de 1 metro golpean la Tierra diariamente, pero causan poco daño. En el otro extremo, objetos muy grandes que causan una extinción masiva sólo chocan una vez cada cien millones de años.

Los casos intermedios quizá son los más preocupantes. La gráfica indica que los objetos que causarían “amplia devastación” —tal como acabar con la población de una ciudad grande— puede esperarse con una frecuencia de una vez cada mil años. Es tan frecuente que justifica al menos algunas acciones preventivas. En la actualidad los astrónomos tratan de hacer predicciones más precisas acerca de cuándo un objeto podría chocar con la Tierra. Si descubren un objeto que chocaría con la Tierra, necesitarían encontrar una manera para desviarlo y prevenir el impacto.

Gráficas con cambio porcentual

Primero obtenga su información y luego puede distorsionarla tanto como quiera.

—Mark Twain

¿Ir a la universidad se está volviendo más o menos caro? Una mirada rápida a la figura 3.44a podría dar la impresión que el costo para las universidades privadas se ha mantenido estable mientras que el costo para universidades públicas disminuyó de manera abrupta en 2006, después de estar subiendo en los años anteriores. Pero vea con más cuidado y verá que éste no es el caso. El eje vertical de las x en esta gráfica representa el porcentaje de incremento en los costos. Una gráfica plana sólo significa que los costos aumentaron el mismo porcentaje cada año, no que los costos se mantuviesen estables. De manera similar, la caída en 2006 para universidades públicas sólo significa que el costo se elevó menos que en los años precedentes.

En realidad, los costos reales (no ajustados por la inflación) para universidades públicas y privadas se han elevado de manera sustancial con el tiempo, como se muestra en la

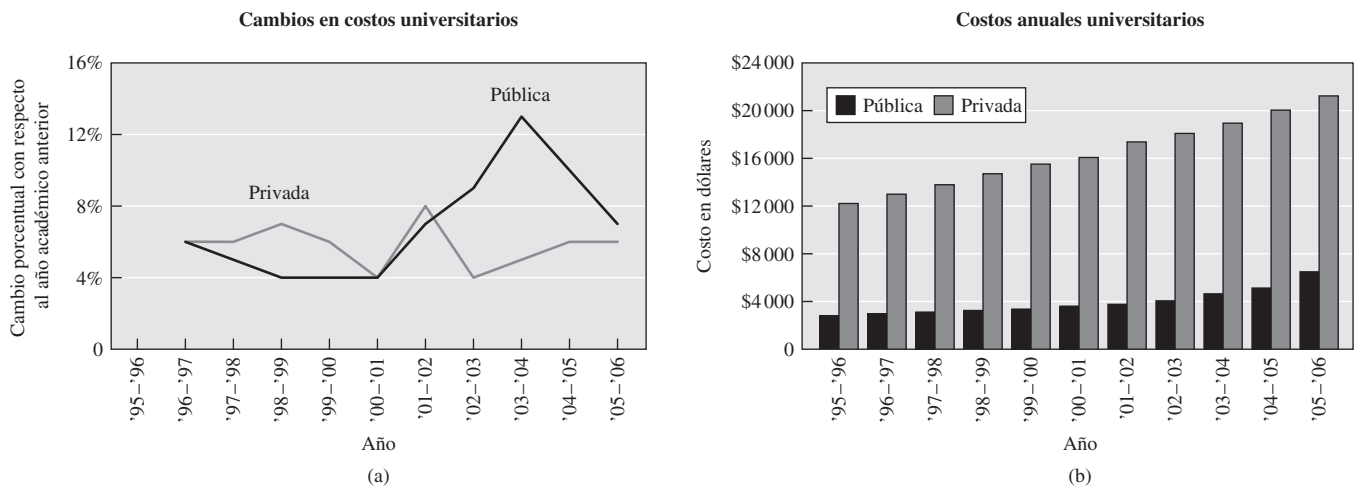


Figura 3.44 Tendencias en costos universitarios: (a) cambio porcentual anual; (b) costos anuales. *Fuente:* The College Board.

figura 3.44b. Además, como la tasa de inflación (medida con el IPC) ha sido menor que la tasa de aumento en los costos universitarios, el costo real de universidades públicas se ha elevado constantemente. Las gráficas que muestran cambios porcentuales son muy comunes, en particular con datos económicos. Aunque son perfectamente honestas, se puede confundir con facilidad, a menos que las interprete con mucho cuidado.

Pictogramas

Los *pictogramas* son gráficas adornadas con trabajo de arte adicional. Este trabajo hace que la gráfica sea más atractiva, pero también puede distraer o confundir. La figura 3.45 es un pictograma que muestra el aumento en la población mundial de 1804 a 2054 (los números para años futuros están basados en las proyecciones de las Naciones Unidas). Las longitudes de las barras de manera correcta corresponden a la población mundial para los años listados. Sin embargo, los adornos artísticos de esta gráfica son engañosos por varias razones. Por ejemplo, su ojo podría desviarse a las figuras de la gente alineada en el globo terráqueo. Puesto que esta línea de personas sube desde el lado izquierdo del pictograma hacia el centro y luego descende, podría

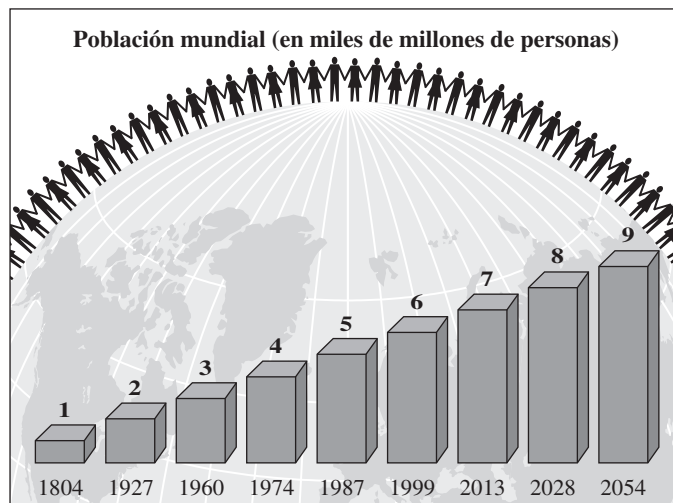


Figura 3.45 Un pictograma. *Fuente:* Datos provenientes de División de Población de Naciones Unidas, Posibilidades de Población Mundial.

A propósito...

Con frecuencia los demógrafos caracterizan el crecimiento de la población por un tiempo de duplicación, el tiempo que tarda la población en duplicarse. Durante el final del siglo xx, el tiempo de duplicación para la población humana era alrededor de 40 años. Si la población continuara duplicándose a esta velocidad, la población mundial llegaría a 34 mil millones en 2100 y 192 mil millones en 2200. Alrededor de 2650 la población humana sería tan grande que no cabría en la Tierra, y todos estarían codo con codo en todas partes.

dar la impresión que la población futura del mundo disminuirá. En realidad la línea de personas es sólo decorativa y no lleva información.

Tal vez el problema más grave con este pictograma es que hace parecer que la población mundial se ha elevado linealmente. Sin embargo, observe que los intervalos de tiempo en el eje horizontal no son los mismos en cada caso. Por ejemplo, el intervalo entre las barras para mil millones y dos mil millones de personas es de 123 años (de 1804 a 1927), pero el intervalo entre las barras para 5 mil millones y 6 mil millones de personas es sólo de 12 años (de 1987 a 1999).

Los pictogramas son muy comunes. Sin embargo, como lo muestra este ejemplo, usted tiene que estudiarlos cuidadosamente para sacar la información esencial y no distraerse por los adornos.

Sección 3.4 Ejercicios

Alfabetización estadística y pensamiento crítico

- Escala vertical.** ¿Por qué es confusa una gráfica cuando su escala vertical no inicia en cero?
- Pictograma.** ¿Qué es un pictograma? Identifique brevemente una ventaja y una desventaja de un pictograma.
- Escala exponencial.** ¿Qué es una escala exponencial? ¿Cuándo es útil usar una escala exponencial en una gráfica?
- Gráficas de cambio porcentual.** Explique cómo una gráfica que muestra cambios porcentuales puede mostrar barras descendentes (o una línea descendente) aunque la variable de interés sea creciente.

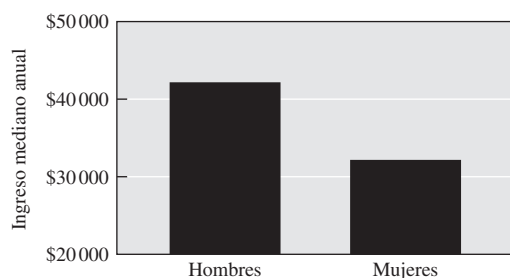


Figura 3.47 Fuente: Oficina del Censo de Estados Unidos.

- Pictograma.** La figura 3.48 muestra las cantidades diarias de consumo de petróleo en Estados Unidos y Japón. ¿La ilustración muestra de manera precisa los datos? ¿Por qué sí o por qué no?

Consumo diario de petróleo
(millones de barriles)

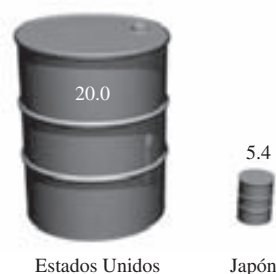


Figura 3.48

Conceptos y aplicaciones

- Distancias de frenado.** La figura 3.46 muestra la distancia de frenado para diferentes automóviles, medida bajo las mismas condiciones. Analice las formas en que la presentación podría ser engañosa. ¿Cuán grande es la distancia de frenado de un Acura RL comparada con la distancia de un Volvo S80? Haga la presentación de manera que muestre los datos de forma más fidedigna.

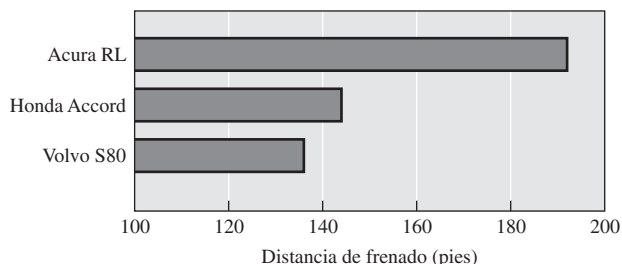


Figura 3.46 Distancias de frenado.

- Comparación de ingresos.** Considere la gráfica de barras en la figura 3.47, la cual compara los ingresos medianos de hombres y mujeres. Identifique cualquier aspecto engañoso de la presentación. Haga la presentación de una manera más justa.

- Pictograma.** Revise la figura 3.48 usada en el ejercicio 7 y construya una gráfica de barras para mostrar los mismos datos en una manera más honesta y objetiva.

- Gráficas circulares en tres dimensiones.** Las gráficas circulares en la figura 3.49 dan un porcentaje de estadounidenses en tres categorías de edad en 1990 y en 2050 (proyectada).

- Considere la distribución de edades en 1990. Los porcentajes reales para las tres categorías para 1990 fueron 87.5% (otros), 11.3% (60-84) y 1.2% (85+). ¿La grá-

fica circular muestra estos valores de manera precisa? Explique.

- b. Considere la distribución de edades para 2050. Los porcentajes reales para las tres categorías para 2050 fueron 80.0% (otros), 15.4% (60-84) y 4.6% (85+). ¿La gráfica circular muestra estos valores de manera precisa? Explique.
- c. Usando los porcentajes reales, dados en los incisos a y b, dibuje una gráfica circular plana (en dos dimensiones) para mostrar estos datos. Explique por qué estas gráficas circulares dan una imagen más precisa que las gráficas en tres dimensiones.
- d. Comente sobre las tendencias generales mostradas en las dos gráficas circulares.

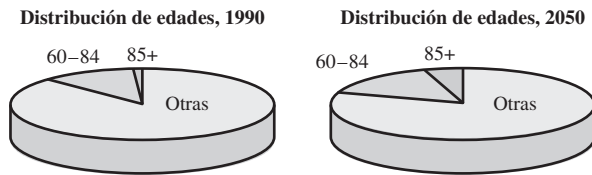
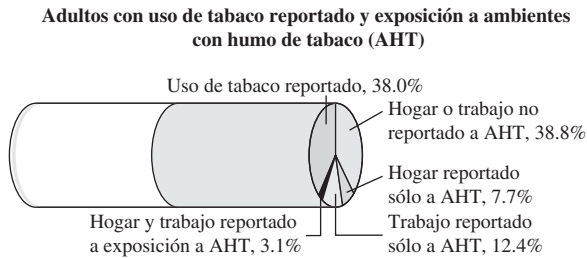


Figura 3.49 Fuente: Oficina del Censo, Estados Unidos.

10. **Pictograma de un cigarro.** Considere el pictograma de la figura 3.50. Describa brevemente qué muestra la gráfica. Analice si es efectiva en su propósito y si es engañosa de alguna manera.



Fuente: Centros de Control y Prevención de Enfermedades

Figura 3.50 Fuente: Wall Street Journal Almanac.

11. **Suscripciones a teléfonos celulares.** La tabla siguiente muestra los números de suscripciones a teléfonos celulares en Estados Unidos.

Año	Número	Año	Número
1985	340	1997	55 312
1987	1231	1999	86 047
1989	3 509	2001	128 375
1991	7 557	2003	158 722
1993	16 009	2005	207 900
1995	33 786		

- a. Construya una gráfica de serie de tiempo de estos datos, usando una escala uniforme en ambos ejes.
- b. Construya una gráfica exponencial de estos datos, en los que las subdivisiones en el eje vertical son 1, 10, 100, 1 000, 10 000, 100 000 y 1 000 000.
- c. Compare las gráficas de los incisos a y b.

12. **Ley de Moore.** En 1965 el cofundador de Intel, Gordon Moore, inició lo que desde entonces se conoció como *ley de Moore*: el número de transistores por pulgada cuadrada en los circuitos integrados se duplicará aproximadamente cada 18 meses. La tabla siguiente lista el número de transistores (en miles) para diferentes años.

Año	Transistores	Año	Transistores
1971	2.3	1993	3100
1974	5	1997	7500
1978	29	1999	24 000
1982	120	2000	42 000
1985	275	2002	220 000
1989	1180	2003	410 000

- a. Construya una gráfica de serie de tiempo de estos datos, use una escala uniforme en ambos ejes.
- b. Construya una gráfica exponencial de estos datos, en los que las subdivisiones en el eje vertical son 1, 10, 100, 1 000, 10 000, 100 000 y 1 000 000.
- c. Compare las gráficas en los incisos a y b.

13. **Cambio porcentual en el IPC.** La gráfica en la figura 3.51 muestra el cambio porcentual en el IPC durante 17 años. ¿En qué año (de los años que se muestran) fue mayor el cambio del IPC? ¿En qué año fue menor el cambio en el IPC? ¿Cómo se comparan los precios en 1991 con los de 1990? Con base en esta gráfica, ¿qué puede concluir acerca de los cambios en los precios durante el periodo que se muestra?

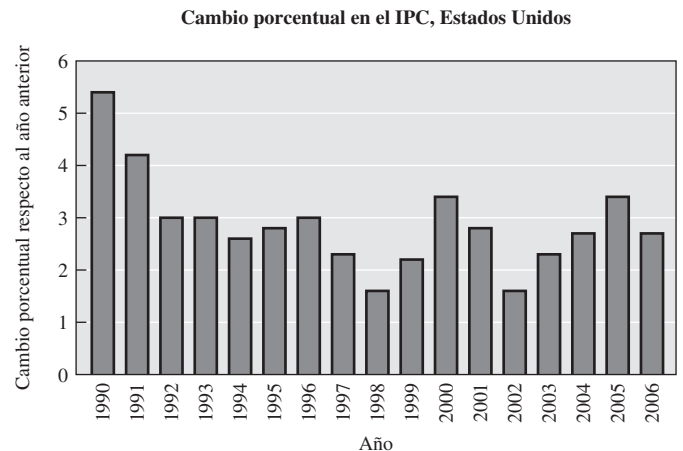


Figura 3.51 Fuente: Oficina de Estadísticas del Trabajo, Estados Unidos.

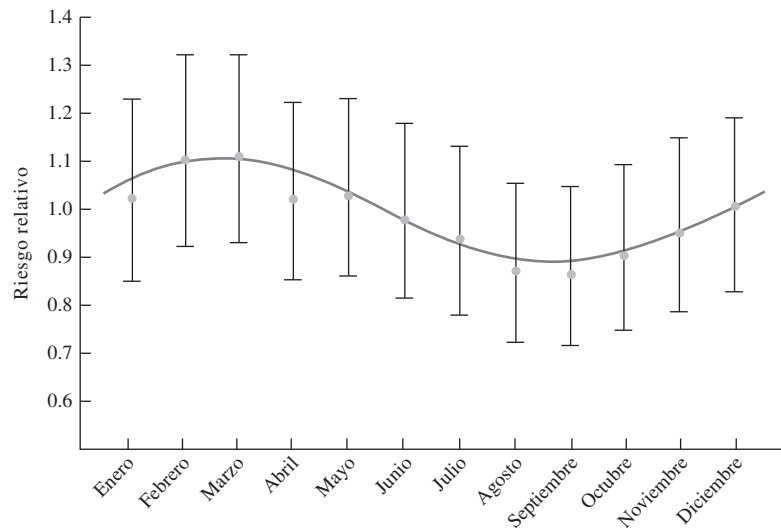


Figura 3.52 Fuente: *New England Journal of Medicine*.

14. ¿Efectos estacionales de la esquizofrenia? La gráfica en la figura 3.52 muestra datos que consideran el riesgo relativo de esquizofrenia en la gente nacida en diferentes meses.

- Observe que la escala del eje vertical no incluye el cero. Haga un bosquejo de la misma curva de riesgo usando un eje que incluya al cero. Comente sobre el efecto de este cambio.
- Cada valor de riesgo relativo se muestra con un punto con su valor más probable y con una “banda de error” que indica el rango en el que probablemente esté el dato. El estudio concluye que “el riesgo también estuvo asociado de manera significativa con la estación de nacimiento”. Dado el tamaño de las barras de error, ¿parece que está justificada la afirmación? (¿Es posible dibujar una línea horizontal que pase por todas las barras de error?)

15. Dólares constantes. La gráfica en la figura 3.53 muestra el salario mínimo en Estados Unidos, junto con su poder de compra

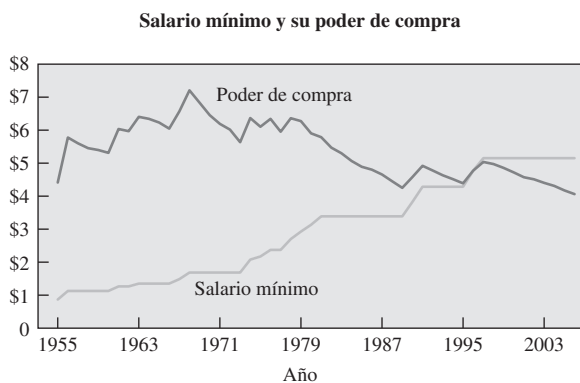


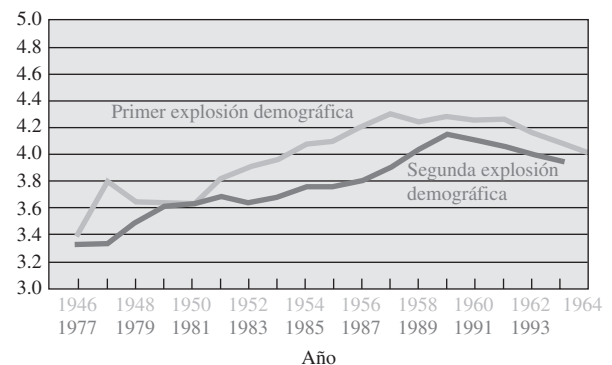
Figura 3.53 Fuente: Departamento del Trabajo, Estados Unidos.

compra, que está ajustado por la inflación con 1996, usado como el año de referencia. La gráfica representa los años de 1955 a 2006. Resuma lo que muestra la gráfica.

16. Doble escala horizontal. La gráfica en la figura 3.54 muestra de manera simultánea el número de nacimientos en Estados Unidos durante dos periodos: 1946-1964 y 1977-1994. ¿Cuándo ocurrió el pico de la primera explosión demográfica? ¿Cuándo ocurrió el pico de la segunda explosión demográfica? ¿Por qué cree que los diseñadores de esta presentación eligieron superponer los dos intervalos de tiempo, en lugar de utilizar una sola escala de 1946 a 1994?

Baby Boomers y sus hijos

La generación de la explosión demográfica (baby boomers), nacidos entre 1946 y 1964, produjeron su propia pequeña explosión demográfica entre 1977 y 1994.



Fuente: Centro Nacional para Estadísticas de Salud

Figura 3.54 Fuente: *Wall Street Journal Almanac*.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 3 en www.aw.com/bbt.

17. Estados Unidos en imágenes. La revista *USA Today* ofrece un pictograma diario para su sección “USA Snapshot”. Obtenga una fotografía de los números de esta semana de *USA Today*. Analice brevemente su propósito y efectividad.

18. Creación de presentaciones. Actualmente las mujeres ganan menos que los hombres haciendo el mismo trabajo. Determine el monto actual ganado por mujeres por cada \$1 ganado por hombres. Dibuje una gráfica que presente esta información de manera objetiva y luego dibuje otra gráfica que exagere la diferencia.

EN LAS NOTICIAS

19. Distorsiones en las noticias. Encuentre un ejemplo del reporte de una gráfica en una noticia reciente que involucre algún tipo de distorsión sensorial. Explique los efectos de la distorsión y describa cómo la gráfica podría haber sido dibujada de manera más honesta.

20. Problemas de escalas en las noticias. Encuentre un ejemplo del reporte de una gráfica en una noticia reciente en la cual la escala vertical no inicie en cero. Sugiera por qué la gráfica fue dibujada de esa manera y también analice las formas en que la gráfica podría ser engañosa a consecuencia de esto.

21. Gráficas económicas en las noticias. Encuentre un ejemplo del reporte de una gráfica en una noticia reciente que muestre datos económicos en el tiempo. ¿Los datos están ajustados por la inflación? Analice el significado de la gráfica y algunas formas en las que podría ser engañosa.

22. Pictogramas en las noticias. Encuentre un ejemplo de un pictograma en una noticia reciente. Analice lo que el pictograma intenta mostrar y si los adornos ayudan o entorpecen este propósito.

23. Gráficas sobresalientes en noticias. Encuentre una gráfica de una noticia reciente que, en su opinión, sea verdaderamente excepcional en la exhibición visual de los datos. Analice lo que la gráfica muestra y explique por qué considera que es sobresaliente.

24. Gráficas no tan sobresalientes. Encuentre una gráfica de una noticia reciente que, en su opinión, falle en su intento de mostrar datos visualmente de una manera significativa. Analice lo que la gráfica está tratando de mostrar, explique por qué falló y cómo podría haber sido mejor.

Ejercicios de repaso del capítulo

A continuación se miden los pesos (en libras) de los contenidos en muestras de latas de Pepsi regular y de Pepsi de dieta. Utilice estos datos para los ejercicios 1-3.

Pepsi regular:

0.8258	0.8156	0.8211	0.8170	0.8216	0.8302
0.8192	0.8192	0.8271	0.8251	0.8227	0.8256
0.8139	0.8260	0.8227	0.8388	0.8260	0.8317
0.8247	0.8200	0.8172	0.8227	0.8244	0.8244
0.8319	0.8247	0.8214	0.8291	0.8227	0.8211
0.8401	0.8233	0.8291	0.8172	0.8233	0.8211

Pepsi de dieta:

0.7925	0.7868	0.7846	0.7938	0.7861	0.7844
0.7795	0.7883	0.7879	0.7850	0.7899	0.7877
0.7852	0.7756	0.7837	0.7879	0.7839	0.7817
0.7822	0.7742	0.7833	0.7835	0.7855	0.7859
0.7775	0.7833	0.7835	0.7826	0.7815	0.7791
0.7866	0.7855	0.7848	0.7806	0.7773	0.7775

1. a. Construya una tabla de frecuencias para los pesos de Pepsi regular. Utilice las clases

0.8130–0.8179
0.8180–0.8229
0.8230–0.8279
0.8280–0.8329
0.8330–0.8379
0.8380–0.8429

- b. Construya una tabla de frecuencias para los pesos de Pepsi de dieta. Utilice las clases

0.7740–0.7779
0.7780–0.7819
0.7820–0.7859
0.7860–0.7899
0.7900–0.7939

- c. Compare las tablas de frecuencias de los incisos a y b. ¿Qué diferencias notables hay? ¿Cómo pueden explicarse esas diferencias?

2. a. Construya una tabla de frecuencias relativas para los pesos de Pepsi regular. Utilice las clases

0.8130–0.8179
0.8180–0.8229
0.8230–0.8279
0.8280–0.8329
0.8330–0.8379
0.8380–0.8429

- b. Construya una tabla de frecuencias acumuladas para los pesos de Pepsi regular.

3. a. Utilice el resultado del ejercicio 1a para construir un histograma de los pesos de Pepsi regular.

- b. Utilice el resultado del ejercicio 1b para construir un histograma para los pesos de Pepsi de dieta.

- c. Compare los histogramas de los incisos a y b. ¿En qué son parecidos y en qué son diferentes?

4. **Gráfica circular.** En un año reciente, 5 524 personas murieron mientras trabajaban. A continuación un desglose de las causas:

transportación: 2 375

contacto con objetos o equipo, 884

asalto o actos de violencia, 829

caídas, 718

exposición a sustancias dañinas o un ambiente dañino, 552

fuego o explosiones, 166

(Los datos son de la Oficina de Estadísticas del Trabajo).

Construya una gráfica circular que represente los datos dados.

5. **Diagrama de Pareto.** Construya un diagrama de Pareto de los datos dados en el ejercicio 4. Compare el diagrama de Pareto con la gráfica circular. ¿Cuál gráfica es más efectiva para mostrar las causas de muerte en el trabajo? Explique.

6. **Gráfica de barras.** La figura 3.55 muestra los números de adopciones en Estados Unidos de China en los años 2005 y 2000. ¿Qué es incorrecto en esta gráfica? Dibuje una gráfica que muestre los datos de una manera más objetiva.

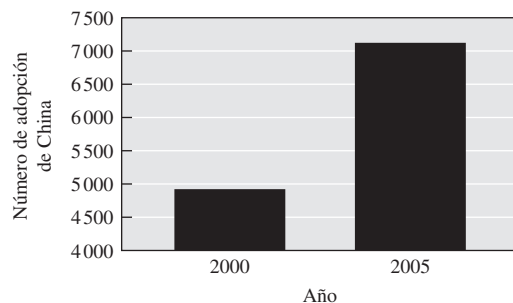


Figura 3.55

Cuestionario del capítulo

1. Las frecuencias de diferentes colores de ojos se obtuvieron de una muestra de sujetos seleccionados aleatoriamente, quienes participaron en un estudio de salud nacional. ¿Cuál gráfica sería más adecuada para estos datos: un histograma, una gráfica de barras, una gráfica de barras múltiple o una gráfica apilada?
2. Para investigar el género y el color de ojos, las frecuencias de diferentes colores de ojos se obtuvieron de una muestra aleatoria de hombres y otra muestra aleatoria de mujeres quienes participaron en un estudio de salud nacional. ¿Cuál gráfica sería más adecuada para estos datos: un histograma, una gráfica de barras, una gráfica de barras múltiple o una gráfica apilada?
3. Se mide el ancho de caderas de sujetos elegidos aleatoriamente, quienes participaron en un estudio de salud nacional. ¿Cuál gráfica sería más adecuada para estos datos: un histograma, una gráfica de barras, una gráfica de barras múltiple o una gráfica apilada?
4. La primera clase en una tabla de frecuencia es 50-59 y la frecuencia correspondiente es 7. ¿Qué indica el valor de 7?
5. La primera clase en una tabla de frecuencia relativa es 50-59 y la frecuencia relativa correspondiente es 0.2. ¿Qué indica el valor de 0.2?
6. Identifique los valores representados en el siguiente diagrama de tallo y hojas.

0	5
0	
0	
1	0
1	22
1	4

7. La gráfica de barra en la figura 3.56 muestra el número de nacimientos de gemelos en Estados Unidos en los años 2000 y 2004. ¿De qué manera es confusa esta gráfica?

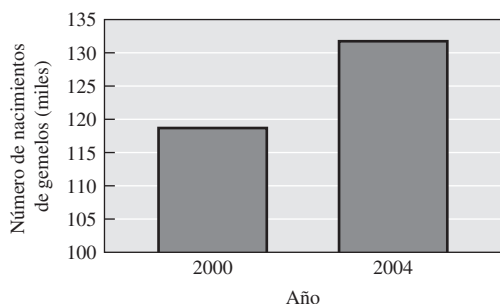


Figura 3.56

8. Identifique el valor más grande representado en la gráfica de puntos en la figura 3.57.

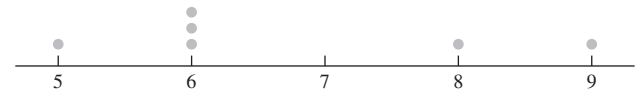
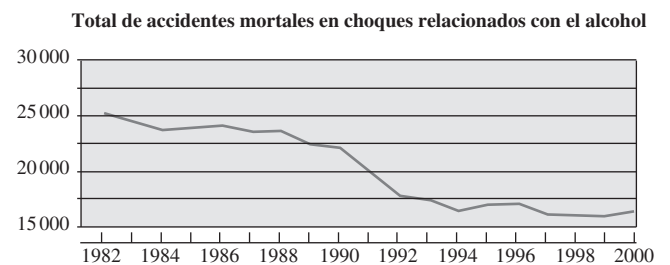


Figura 3.57

9. La figura 3.58 muestra el número de accidentes automovilísticos mortales en Estados Unidos en los que el alcohol estuvo involucrado para cada año desde 1982 a 2000. ¿Cuántos accidentes mortales hubo en 1982? ¿Cuántos en 2000?



Fuente: Departamento Nacional de Seguridad en Carreteras

Figura 3.58

10. Un reportero de un periódico quiere construir una gráfica que muestre la cantidad de gasolina consumida anualmente por automóviles en Estados Unidos para una sucesión de años recientes. ¿Por qué no tiene sentido ajustar las cantidades anuales usando un año de referencia particular, tal como 1996?

Uso de tecnología

Los paquetes de cómputo ahora son muy efectivos para la generación de una amplia variedad de gráficas impresionantes. Las instrucciones detalladas pueden variar de muy fáciles a muy complejas, así que a continuación proporcionamos comentarios pertinentes. Procedimientos a mayor detalle pueden encontrarse en el *Technology Manual and Workbook*. Además, estos complementos de la serie de Estadística de Triola, todos disponibles de Pearson Addison Wesley, pueden ser útiles: *SPSS Student Laboratory Manual and Workbook* de James J. Ball, *Excel Student Laboratory Manual and Workbook* de Johana Halsey y Ellena Reda, y *Statdisk Student Laboratory Manual and Workbook* por Mario F. Triola. Los complementos también están disponibles para Minitab, SAS y la calculadora TI-83/84.

SPSS

SPSS puede usarse para crear gráficas de barras, gráficas de líneas, gráficas circulares, diagrama de Pareto, histogramas, y gráficas de series de tiempo. Primero entre o abra el conjunto de datos deseado. Luego de clic en **Graphs** y proceda a seleccionar el tipo deseado de gráfica. Por ejemplo, para generar un histograma seleccione el elemento del menú **Histogram**, de clic en la columna de datos deseado, de clic en botón siguiente en la etiqueta **Variable**, y luego de clic en **OK**. Nota: la versión estudiantil de SPSS está limitada a no más de 1500 valores de datos.

Excel

Aunque puede generar histogramas, gráficas de barras, gráficas de líneas, gráficas circulares, gráficas apiladas, gráficas de líneas apiladas y gráficas de múltiples barras, por lo regular Excel es difícil de utilizar. El complemento DDXL puede instalarse en Excel, y puede usarse para generar fácilmente un histograma, una gráfica de barras, una gráfica de puntos o una gráfica circular. El complemento DDXL está disponible en www.aw.com/bbt. Después de que se ha instalado DDXL en Excel, proceda como sigue.

Primero introduzca los datos en la columna A, luego dé clic en el complemento **DDXL**, seleccione **Charts and Plots**, dé clic en el cuadro etiquetado **function type** y del menú seleccione la gráfica deseada. Por ejemplo, para generar un histograma seleccione **Histogram** del menú, luego dé clic en el icono del lápiz e ingrese el rango de celdas que contiene los datos, tal como A1:A500 para 500 valores en los renglones 1 a 500 de la columna A.

STATDISK

STATDISK es gratis para quienes compran copias nuevas de este libro, y está disponible en www.aw.com/bbt. STATDISK genera con facilidad histogramas. Ingrese los datos en la ventana de datos de STATDISK, dé clic en **Data**, dé clic en **Histogram** y luego dé clic en el botón **Plot**. (Si prefiere introducir su propio ancho de clase y el punto inicial, dé clic en el botón **User defined** antes de dar clic en **Plot**).

Para generar una gráfica circular, primero ingrese los nombres de las categorías en la columna 1 de la ventana de datos de STATDISK y luego ingrese las frecuencias correspondientes en la columna 2; seleccione **Data**, dé clic en **Pie Chart** y dé clic en **Chart**.

¿La guerra puede describirse con una gráfica?

¿Una gráfica puede describir la guerra? La figura 3.59 (en la página 138) creada por Charles Joseph Minard en 1889, lo hace extraordinariamente bien. Esta gráfica cuenta la historia de la desafortunada campaña de 1812 de Napoleón, algunos le han llamado la marcha de la muerte de Napoleón.

El mapa subyacente en la gráfica de Minard muestra una franja de aproximadamente 500 millas de tierra que va desde el río Niemen, en la frontera polaco-rusa hasta Moscú. La primera franja presenta la marcha hacia adelante del ejército de Napoleón. En el dibujo original de Minard cada milímetro de anchura representaba 6 000 hombres, esta reproducción se muestra en un tamaño más pequeño que el original. La marcha comienza en la izquierda lejana donde la franja es más ancha. Aquí, un ejército de 422 000 hombres comienza a marchar triunfalmente hacia Moscú el 24 de junio de 1812. En ese momento, era el ejército más grande jamás movilizado.

La reducción de la franja en cuanto se aproxima a Moscú representa cómo se diezmaba el ejército. (Las ramificaciones representan los batallones que fueron enviados en otras direcciones a lo largo del camino). Napoleón sólo había llevado el mínimo de víveres y el caluroso verano acompañado de fuertes lluvias trajeron incontrolables enfermedades. Inanición, enfermedades y miles de hombres muertos en combate cada día. Para cuando el ejército llegó a Moscú el 14 de septiembre, se había reducido a 100 000 hombres. Lo peor estaba por venir.

Para consternación de Napoleón, los rusos evacuaron Moscú antes de la llegada del ejército francés. Carente de oportunidad para atrapar a las tropas rusas sabiendo que las condiciones de su ejército eran pésimas para continuar hacia la capital rusa de San Petersburgo, Napoleón tomó sus tropas y las llevó rumbo al sur en las afueras de Moscú. La parte más baja de la franja sobre la gráfica (la segunda línea) representa la retirada, y la última línea abajo marca la temperatura que tenían por las noches con el invierno aproximándose. Las heladas temperaturas ya estaban desde el 18 de octubre.

Las temperaturas caían bajo 0°F a finales de noviembre. La repentina angostura de la franja inferior, alrededor del 28 de noviembre, muestra que 22 000 hombres perecieron en los bancos del río Berezina. Tres cuartas partes de los sobrevivientes se congelaron hasta morir durante los siguientes días, muchos de ellos por el crudo frío nocturno del 6 de diciembre. Para cuando el ejército llegó a Polonia el 14 de diciembre, sólo lo hizo con 10 000 de los 422 000 hombres iniciales.

En un famoso análisis de técnicas gráficas, el autor Edward Tufte describió la gráfica de Minard como la posible “mejor gráfica jamás dibujada”. Sin embargo, E. J. Marey, un contemporáneo de Minard, escribió que esta gráfica “causó lágrimas en los ojos de toda Francia”.



PREGUNTAS PARA DISCUSIÓN

1. Discuta cómo esta gráfica ayuda a vencer la naturaleza impersonal de muchas muertes en una guerra. ¿Qué impacto le causó a usted personalmente?
2. Observe que esta gráfica traza seis variables: dos variables de dirección (norte-sur y este-oeste), el tamaño del ejército, la dirección de movilización del ejército y las temperaturas durante la retirada. ¿Considera usted que Minard pudo haber conseguido los puntos de coincidencia con menos variables? ¿Por qué sí o por qué no?
3. Discuta cómo se puede hacer una gráfica para algún otro evento histórico o político.

LECTURAS SUGERIDAS

Tufte, Edgard, *The Visual Display of Quantitative Information*, Graphic Press, 1992.

Wainer, Howard, *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*, Copernicus, Nueva York, 1997.

HABLEMOS DE MEDIO AMBIENTE



¿Cómo podemos visualizar el calentamiento global?

El calentamiento global es sin duda uno de los temas más discutidos de nuestro tiempo. Pero, ¿cuán serio es el problema? y ¿cómo podemos decidir si el calentamiento es causado por la actividad humana o por los ciclos naturales de la Tierra? Estas preguntas pueden ser abordadas estadísticamente y, como muchas otras ideas estadísticas, se visualizan mucho mejor a través de gráficas.

El punto obvio de partida para un estudio científico sobre el calentamiento global es averiguar qué tanto calentamiento está ocurriendo en realidad. Lo primero que se puede hacer para estudiar tendencias del clima es simplemente comparar las temperaturas locales en periódicos antiguos con las temperaturas actuales de los mismos lugares. Sin embargo, el calentamiento global se refiere al incremento en el *promedio* de temperatura de todo el planeta, lo que significa que localidades en particular puedan calentarse más o menos que este promedio. De hecho, deberíamos esperar que algunos lugares pudieran enfriarse aun cuando el planeta sufre calentamiento. Si queremos saber si el planeta se calienta entonces necesitamos información que nos diga cómo la *temperatura promedio global* —la temperatura promedio de la Tierra— está cambiando con el tiempo.

En la actualidad, los satélites que están en órbita proporcionan información que nos permite determinar el promedio de temperatura con mucha precisión, ya que nos proporcionan una vista general del planeta. Podemos validar estos registros con mediciones “verdaderas en tierra” registradas en más de 7000 estaciones meteorológicas alrededor del mundo, junto con medidas de temperaturas del océano, generalmente obtenidas con la medición de la temperatura del agua recolectada por las válvulas de admisión de los barcos. Como resultado, tenemos información fidedigna de temperatura para las casi cuatro décadas para las que tenemos observaciones satelitales de nuestro planeta. La información que proviene de algunos años antes de la era satelital es un poco menos confiable porque sólo podemos ver información de localidades específicas, y el número de localidades que podemos ver en los dos siglos anteriores es más reducida. Además, la mayoría de los datos de las temperaturas capturadas fueron de ciudades, las cuales tienden a ser más calientes con respecto al tiempo por razones independientes del calentamiento global (vea el ejemplo 1 de la sección 2.2). Con frecuencia los científicos explican este efecto de “islas urbanas de calor”, pero incluso entonces se dejan principalmente datos para temperaturas en tierra y pocos registros de temperaturas sobre los océanos, los cuales cubren tres cuartas partes de la superficie terrestre. Los científicos han dedicado mucho de su esfuerzo, de los últimos años, a examinar información de la temperatura del pasado a detalle. A través de análisis estadísticos es posible reconstruir la historia de la temperatura justa y fidedigna de la mayor parte de los dos siglos anteriores, aunque la incertidumbre se hace mayor conforme vemos más atrás.

La figura 3.60 muestra la historia reconstruida de la temperatura promedio global de la Tierra desde 1860. A pesar de ciertas dudas, la conclusión general es clara: la temperatura promedio global se ha elevado en alrededor de 0.8°C (1.4°F) en el siglo pasado. Además, la mayoría del calentamiento (cerca de 0.6°C) ocurrió en sólo los últimos treinta años, el periodo cuya información es la más confiable. El calentamiento al parecer tiende a acelerarse. Por ejemplo, nueve de los diez años más calurosos registrados ocurrieron en el periodo de diez años más recientes, mostrado en la gráfica.

Con la información de la temperatura que muestra una clara tendencia al calentamiento, la siguiente pregunta es ¿cuánto podríamos esperar que la temperatura se incremente en el futuro?

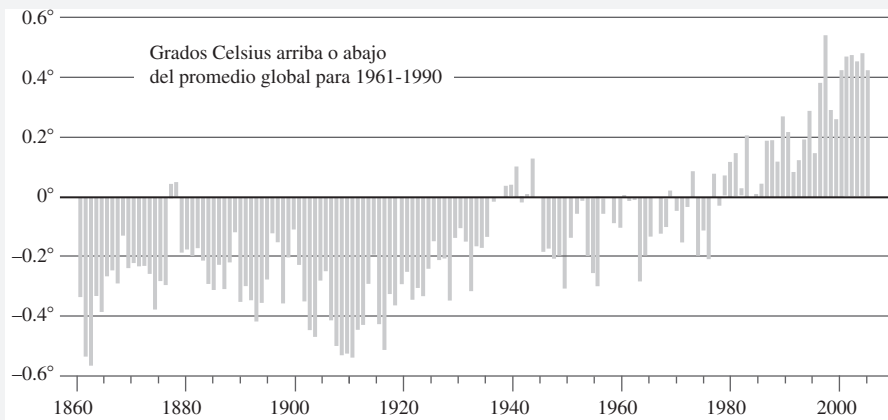


Figura 3.60 Temperaturas globales promedio desde 1860, comparadas con la temperatura promedio durante el periodo de 30 años de 1961 a 1990. Observe la clara tendencia al calentamiento global de las pasadas décadas. Fuente: Centro Nacional de Investigación Atmosférica.

Como lo analizaremos más adelante en el capítulo 7 (vea la sección “Hablemos de” en la página 328), la elevación de la temperatura claramente está asociada a un incremento en la concentración atmosférica de dióxido de carbono (CO_2) y otros *gases de invernadero*. Por tanto, si queremos predecir cuánto se elevará la temperatura en el futuro, necesitamos datos sobre la reciente elevación en la concentración atmosférica de CO_2 , de igual manera se requiere la información anterior referente a cómo la concentración de CO_2 está asociada a los cambios climáticos globales.

Para el reciente aumento en la concentración de CO_2 tenemos una extraordinaria información obtenida desde 1958 por científicos que laboran en una estación de Mauna Loa (Hawái). Estos datos mostrados del lado derecho de la figura 3.61 representan medidas directas del cambio de concentración de CO_2 . Note que las unidades de la concentración son *partes por millón* (ppm), por ejemplo: 320 ppm significa que hay 320 moléculas de dióxido de carbono en cada millón de moléculas de aire. Las oscilaciones anuales en la gráfica muestran que la concentración del dióxido de carbono varía con las estaciones. A pesar de estas oscilaciones, la tendencia es claramente hacia arriba.

Para saber cuánto podríamos esperar que la temperatura suba, nos gustaría tener información que muestre la correlación entre la concentración del dióxido de carbono y la temperatura

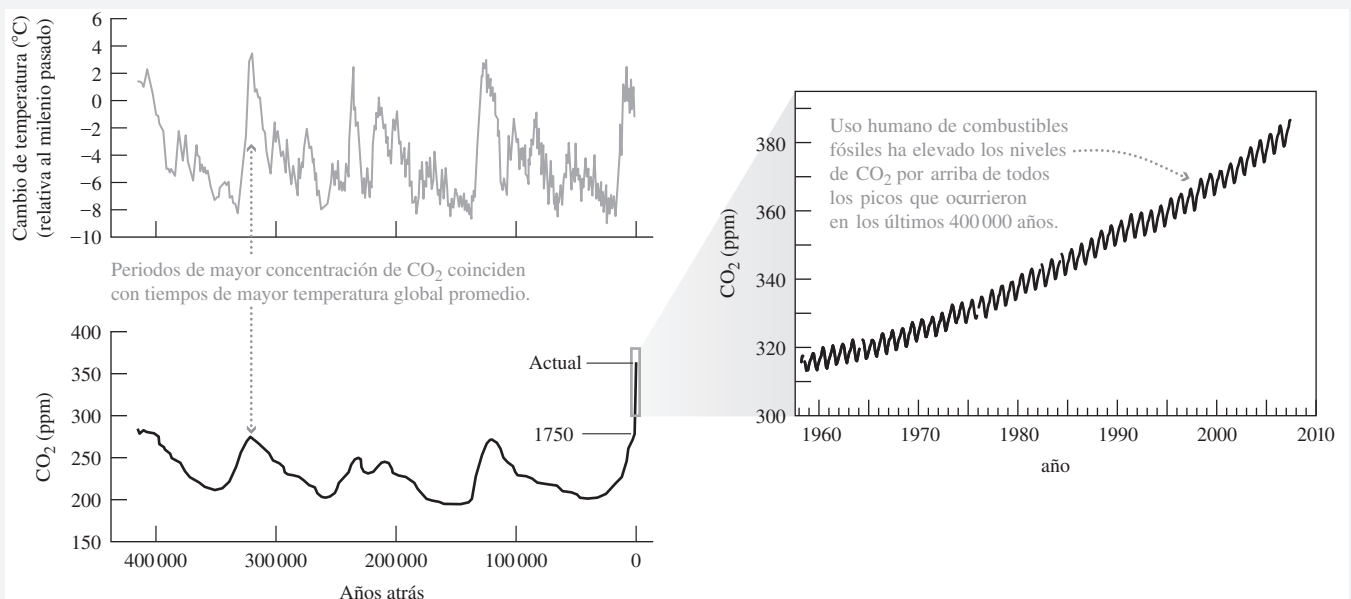


Figura 3.61 La concentración atmosférica de dióxido de carbono y la temperatura global promedio reconstruidas con base en núcleos de glaciares y otros datos para los últimos 400 000 años. Los datos recientes de CO_2 (derecha) representan mediciones directas tomadas en Mauna Loa (Hawái).

en el pasado. Para los últimos milenios los científicos pueden obtener esta información examinando los anillos de los árboles; con un estudio cuidadoso pueden revelar un registro tanto de temperatura como de concentración de dióxido de carbono. Los registros pueden extenderse mucho más atrás con estudios de los núcleos de hielo extraído de capas heladas de la Antártica. Los núcleos de hielo se formaron de capas antiguas acumuladas, comprimiendo nieve. Ya que la nieve cae cada temporada, cada capa delgada representa un año, así como cada anillo de un árbol representa un año en la vida de éste. Mediante el estudio de burbujas de aire atrapadas en las capas de hielo, los científicos han podido reconstruir la historia de la temperatura y de la concentración de dióxido de carbono. (La concentración del dióxido de carbono puede medirse directamente de las burbujas de aire mientras las temperaturas se miden a través de estudios cuidadosos de isótopos de oxígeno en las burbujas de aire: los isótopos más pesados no son transportados en el aire tan fácilmente cuando las temperaturas son más bajas, así que la proporción de los isótopos más densos y más ligeros permite a los investigadores estimar temperaturas en el pasado).

Los resultados son sorprendentes. Como puede ver en la gráfica inferior en el lado izquierdo de la figura 3.61 los datos de los núcleos de hielo proporcionan un registro de temperatura y concentración de dióxido de carbono que se remontan a más de 400 000 años, periodo en el cual la Tierra pasó por numerosas edades de hielo así como periodos de calentamiento. Al menos tres conclusiones deben saltarle a la vista cuando estudia la figura:

1. Existe una relación entre temperatura y concentración de dióxido de carbono: los periodos de temperatura más alta tienden a ser también los periodos de mayor concentración de dióxido de carbono. Aunque esto no prueba que la concentración del dióxido de carbono sea la *causa* de la temperatura más alta, hace parecer muy probable que van de la mano —quiere decir que el aumento reciente en la concentración de dióxido de carbono ha sido acompañado por un correspondiente aumento en la temperatura.
2. La temperatura y la concentración del dióxido de carbono varían natural y sustancialmente con el tiempo. La temperatura promedio global ha subido y bajado por más de 10°C (aproximadamente 18°F) en muchas ocasiones en los últimos cientos de miles de años y la concentración del dióxido de carbono ha variado naturalmente entre menos de 200 y casi 300 partes por millón (ppm).
3. Cuando comparamos la información del dióxido de carbono del lado derecho de la gráfica con los datos del pasado del lado izquierdo, encontramos que estamos en un territorio completamente nuevo: la concentración actual de dióxido de carbono está muy por encima de lo que nuestro planeta había visto en los 400 000 años anteriores. En realidad, datos recientes han extendido, de manera preliminar, el registro de núcleos de hielo a un millón de años atrás, y encontramos la misma idea básica. Desde el inicio de la era industrial hemos elevado la concentración atmosférica de dióxido de carbono por arriba de los niveles que han ocurrido naturalmente durante cualquiera de los periodos de calentamiento o eras de hielo que sucedieron millones de años atrás.

Aún sin predicción precisa de los cambios esperados en el futuro sobre el clima global, estos datos proporcionan un mensaje para reflexionar. La diferencia entre una era glacial, en que la mayor parte de Estados Unidos estaría enterrado entre los glaciares y nuestras más actuales condiciones cálidas en sólo algunos grados Celsius, y los datos de los núcleos de hielo nos dicen que incluso grandes cambios de temperatura han ocurrido en el pasado con cambios en la concentración del dióxido de carbono, mucho más pequeños que los cambios que están ocurriendo ahora. Si las tendencias pasadas son indicadores de lo que podemos esperar en el futuro entonces, a menos que hagamos algo pronto para detener el crecimiento del dióxido de carbono, nuestros hijos y nietos habitarán un mundo con un clima muy diferente al que vivimos hoy en día.

PREGUNTAS PARA DISCUSIÓN

1. Estudie cuidadosamente la figura 3.61. ¿Cuánta es la concentración de dióxido de carbono actual comparada con la de 1750? ¿Cómo es esta concentración en el año 1750 comparada con la de hace 400 000 años? ¿Qué conclusiones puede usted sacar de sus respuestas a estas preguntas?

2. Discuta algunos de los factores que pueden afectar la futura concentración atmosférica de dióxido de carbono. ¿Qué debería hacerse para reducir o detener el crecimiento en la concentración del dióxido de carbono?
3. ¿Qué tipo de consecuencias se podrían esperar del calentamiento global? Discuta lo que haya escuchado en las noticias y considere qué consecuencias parecen probables y cuáles parecen excesivamente optimistas o demasiado alarmantes.
4. Hace apenas una década había cierta controversia acerca de si el calentamiento global en realidad estaba ocurriendo o si se podría calificar como un problema serio. Actualmente esa controversia ya está terminada. Discuta el papel que la nueva información ha desempeñado en aclarar la amenaza del calentamiento global. ¿Cuán importantes son gráficas claras para nuestra comprensión de estos datos nuevos?

LECTURAS SUGERIDAS

Usted puede encontrar gran cantidad de sitios científicos en internet con las últimas actualizaciones de la información del calentamiento global. Un buen punto de partida son los sitios del Intergovernmental Panel on Climate Change (IPCC), el Centro Nacional de Investigación Atmosférica (NCAR, por sus siglas en inglés) y la Agencia de Protección Ambiental (EPA, por sus siglas en inglés) de Estados Unidos.

Bennet, Jeffrey y Shostak, Seth, *Life in the Universe*, segunda edición, Addison Wesley, 2007 (véase la sección 10.5).

Flannery, Tim, *The Weather Makers*, Atlantic Monthly Press, 2005.



*Es inútil tratar de hacer caso a todos.
Uno debe seguir sugerencias,
no exactamente como se dicen,
ni completamente como están hechas.*
—Virginia Woolf

Descripción de datos

EN EL CAPÍTULO 3 ESTUDIAMOS LOS MÉTODOS PARA

mostrar la distribución de datos con tablas y gráficas. Ahora estamos listos para estudiar cómo podemos resumir las características de una distribución en términos de unas cuantas propiedades y números. En particular, analizaremos métodos comunes para la descripción del centro, forma y variación de un conjunto de datos. Estos métodos son primordiales para el análisis de datos y, como veremos, tienen aplicaciones en casi cualquier estudio estadístico que usted encuentre en las noticias. Concluiremos el capítulo estudiando algunas sorpresas que en ocasiones se presentan incluso cuando examinamos los datos cuidadosamente.

OBJETIVOS DE APRENDIZAJE

4.1 ¿Qué es un promedio?

Comprender la diferencia entre una media, mediana y moda y cómo cada una es afectada por datos atípicos. Además, entender cómo estos diferentes tipos de “promedios” pueden llevar a confusión y cuándo es apropiado utilizar una media ponderada.

4.2 Formas de las distribuciones

Ser capaz de describir la forma general de una distribución en términos de su número de modas, sesgo y variación.

4.3 Medidas de variación

Comprender e interpretar estas medidas de variación comunes: rango, resumen de cinco números y desviación estándar.

4.4 Paradojas estadísticas

Investigar algunas paradojas comunes que surgen en estadística, por ejemplo, cómo es posible que la mayoría de la gente que falle en una prueba del polígrafo “con 90% de precisión”, en realidad está diciendo la verdad.

4.1 ¿Qué es un promedio?

El término *promedio* aparece con frecuencia en noticias y otros reportes, pero no siempre tiene el mismo significado. Como veremos en esta sección, la definición más apropiada de *promedio* depende de la situación.

Media, mediana y moda

Tabla 4.1 Películas de cinco series de ciencia ficción	
Serie	Número de películas (hasta 2007)
<i>Alien</i>	4
<i>Volver al futuro</i>	3
<i>El planeta de los simios</i>	6
<i>Viaje a las estrellas</i>	10
<i>La guerra de las galaxias</i>	6

La tabla 4.1 muestra el número de películas (original, secuelas o previos) para cada una de las cinco series de ciencia ficción más populares. ¿Cuál es el número promedio de películas en estas series? Una manera de responder es calcular la **media**. (El término formal de *media aritmética* simplemente es referido como *media*). Determinamos la media al dividir el número total de películas entre cinco (ya que hay cinco series en el conjunto de datos):

$$\text{media} = \frac{4 + 3 + 6 + 10 + 6}{5} = \frac{29}{5} = 5.8$$

En otras palabras, hay cinco series que tienen una media de 5.8 películas. Determinamos la media de cualquier conjunto de datos si dividimos la suma de todos ellos entre el número de datos. La media es lo que la mayoría de las personas consideran como el promedio. En esencia, representa el punto de equilibrio para una distribución de datos cuantitativos, como se muestra en la figura 4.1.

También describimos el número promedio de películas mediante el cálculo de la **mediana**, o valor medio, del conjunto de datos. Para determinar la mediana, acomodamos los valores en orden ascendente (o descendente), repitiendo los datos que aparezcan más de una vez. Si el número de valores es impar existe exactamente un valor en la mitad de la lista, y este valor es la mediana. Si el número de valores es par, existen dos valores en la mitad de la lista, y la mediana es el número que está a la mitad de ellos. Al colocar los datos de la tabla 4.1 en orden ascendente se obtiene la lista 3, 4, 6, 6, 10. El número mediano de películas es 6, ya que 6 es el número a mitad de la lista.

La **moda** es el valor o grupo de valores más común en un conjunto de datos. En el caso de las películas, la moda es 6 ya que este valor ocurre dos veces en el conjunto de datos, mientras que los otros valores aparecen una sola vez. Un conjunto de datos puede tener una moda, más de una moda, o no tener moda. En ocasiones la moda muestra un grupo de valores muy cercanos en lugar de un solo valor. La moda se utiliza de manera más común para datos cualitativos que para datos cuantitativos, cuando ni la media ni la mediana pueden usarse con datos cualitativos.

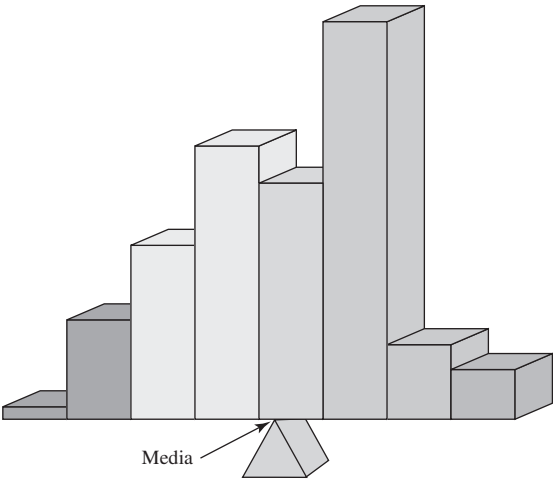


Figura 4.1 Un histograma realizado con bloques que se equilibrarían en la posición de la media.

Definiciones: medidas de tendencia central en una distribución

La **media** es a la que llamamos comúnmente el valor promedio. Se determina como sigue:

$$\text{media} = \frac{\text{suma de todos los valores}}{\text{número total de valores}}$$

La **mediana** es el valor de en medio en el conjunto ordenado de datos (o a la mitad entre los dos valores de la mitad, en caso de que el número de valores sea par).

La **moda** es el valor (o grupo de valores) más común en un conjunto de datos.

Al redondear utilizaremos la regla siguiente para todos los cálculos estudiados en este capítulo.

Regla de redondeo para cálculos estadísticos

Indique su respuesta con *un decimal más* de precisión que los que se encuentran en los datos en procesar. Ejemplo: La media de 2, 3 y 5 es 3.3333..., que redondeamos a 3.3. Puesto que los datos son números enteros, redondeamos al décimo más cercano. *Como siempre, redondee sólo la respuesta final y no los valores intermedios usados en sus cálculos.*

NOTA TÉCNICA

Si la medida de tendencia central tiene el mismo número de dígitos significativos que los datos originales, puede incluir un cero adicional o bien utilizar el resultado exacto sin lugares decimales extra. Por ejemplo, la media de 2 y 4 puede expresarse como 3 o 3.0.

Observe que aplicamos esta regla en nuestro ejemplo de las películas. Los datos en la tabla 4.1 consisten en números enteros, pero indicamos la media como 5.8.

EJEMPLO 1 Datos de precios

Ocho tiendas venden barras de energía PR a los precios siguientes:

\$1.09 \$1.29 \$1.29 \$1.35 \$1.39 \$1.49 \$1.59 \$1.79

Determine la media, la mediana y la moda para estos precios.

Solución El precio *medio* es \$1.41:

$$\begin{aligned}\text{media} &= \frac{\$1.09 + \$1.29 + \$1.29 + \$1.35 + \$1.39 + \$1.49 + \$1.59 + \$1.79}{8} \\ &= \$1.41\end{aligned}$$

Para determinar la *mediana*, primero ordenamos los datos en orden ascendente:

$\underbrace{\$1.09, \$1.29, \$1.29}_{3 \text{ valores abajo}} \quad \underbrace{\$1.35, \$1.39}_{2 \text{ valores en medio}} \quad \underbrace{\$1.49, \$1.59, \$1.79}_{3 \text{ valores arriba}}$

Puesto que hay ocho precios (un número par), existen dos valores a la mitad de la lista: \$1.35 y \$1.39. Por tanto, la mediana está a la mitad entre estos dos valores, que calculamos sumándolos y dividiendo entre 2:

$$\text{mediana} = \frac{\$1.35 + \$1.39}{2} = \$1.37$$

Con la regla de redondeo expresamos la media y la mediana como \$1.410 y \$1.370, respectivamente.

La *moda* es \$1.29 ya que este precio aparece más veces que cualquiera otro.

A propósito...

Un estudio reciente en la Universidad de Chicago descubrió que la asignación mediana para adolescentes en Estados Unidos es \$50 por semana. El estudio estimó que los adolescentes reciben más de mil millones de dólares a la semana en asignaciones.



FRAZZ reimpresso con permiso de United Feature Syndicate, Inc.

EJEMPLO 2 Océanos y mares

La tabla 4.2 lista áreas, en kilómetros cuadrados, de los océanos y mares del mundo. Para estas áreas determine la media, la mediana y la moda.

Tabla 4.2 Áreas de océanos y mares del mundo					
Océano/mar	Área (km²)	Océano/mar	Área (km²)	Océano/mar	Área (km²)
Océano Pacífico	165 760 000	Mar del Sur de China	2 319 000	Mar de Japón	1 007 800
Océano Atlántico	82 400 000	Mar de Bering	2 291 900	Mar Andaman	797 700
Océano Índico	65 527 000	Golfo de México	1 592 800	Mar del Norte	575 200
Océano Ártico	14 090 000	Mar Okhotsk	1 589 700	Mar Rojo	438 000
Mar Mediterráneo	2 965 800	Mar del Este de China	1 249 200	Mar Báltico	422 200
Mar Caribe	2 718 200	Bahía de Hudson	1 232 300		

Fuente: Almanaque de TIME.

Solución Si usted revisa, encontrará que la suma de las áreas de los 17 océanos y mares listados es 346 976 800 kilómetros cuadrados. Determinamos la *media* dividiendo esta suma entre 17:

media = $\frac{346\,976\,800\text{ km}^2}{17} = 20\,410\,400\text{ km}^2$

La tabla 4.2 ya está ordenada en forma descendente, lo cual hace más fácil determinar la mediana. Con 17 elementos en la lista, el noveno valor es la mediana, ya que 8 valores están por arriba de él y 8 están por abajo. El noveno valor en la lista es el Golfo de México, por lo que la *mediana* es el área de 1 592 800 km².

Estos datos no tienen *moda*, ya que ningún valor aparece más de una vez. Sin embargo, observando grupos de valores, 5 de los 17 océanos y mares tienen áreas entre 1 y 2 millones de kilómetros cuadrados. Por tanto, podríamos decir que las áreas más comunes son aquellas entre 1 y 2 millones de kilómetros cuadrados.

Efecto de datos atípicos

Para explorar las diferencias entre la media, la mediana y la moda, imagine cinco estudiantes graduados en un equipo de básquetbol universitario que reciben las siguientes ofertas de

contrato del primer año para jugar en la NBA (cero indica que el jugador no recibió una oferta de contrato):

0 0 0 0 \$3 500 000

La media de oferta de contrato es

$$\text{media} = \frac{0 + 0 + 0 + 0 + \$3\,500\,000}{5} = \$700\,000$$

Por tanto, ¿es correcto decir que el estudiante graduado *promedio* en este equipo de básquetbol recibió una oferta de contrato de \$700 000?

En realidad no. El problema es que un solo jugador recibió la oferta mayor que hizo la media mucho más grande a lo que hubiese sido de otra forma. Si ignoramos a este único jugador y sólo vemos a los otros cuatro, la oferta de contrato media es cero. Ya que este único valor, \$3 500 000, es tan extremo comparado con los otros, decimos que es un **dato atípico**. Como lo muestra nuestro ejemplo, un dato atípico puede jalar a la medida significativamente hacia arriba (o hacia abajo), de este modo haciendo a la media poco representativa del conjunto de datos.

Definición

Un **dato atípico** en un conjunto de datos es un valor que es mucho mayor o mucho menor que la mayoría de los demás valores.

Aunque el dato atípico jala la oferta media de contrato hacia arriba, no tiene efecto sobre la oferta mediana de contrato, la cual permanece en cero para los cinco jugadores. En general, el valor de un dato atípico no tiene efecto sobre la mediana, ya que los valores atípicos no están a la mitad de un conjunto de datos. Los valores atípicos no afectan a la moda. (Sin embargo, la mediana podría cambiar si eliminamos un dato atípico, ya que estamos cambiando el número de valores en el conjunto de datos). La tabla 4.3 resume las características de la media, mediana y moda, incluyendo los efectos de los valores atípicos en cada medida.

UN MOMENTO DE REFLEXIÓN

¿Es correcto utilizar la mediana como la oferta promedio de contrato para los cinco jugadores? ¿Por qué sí o por qué no?

A propósito...

Una encuesta determinó que los graduados en geografía de la Universidad de Carolina del Norte tenían un salario medio inicial más alto que los graduados en geografía de otras universidades. La razón para la media tan alta fue causada por un solo dato atípico: la súper estrella del básquetbol y graduado en geografía Michael Jordan.



Tabla 4.3 Comparación de media, mediana y moda

Medida	Definición	¿Qué tan común?	Existencia	¿Toma todos los valores en cuenta?	¿Afectada por valores atípicos?	Ventajas
Media	suma de todos los valores número total de valores	el “promedio” más familiar	siempre existe	sí	sí	se entiende, funciona bien con muchos métodos estadísticos
Mediana	valor de en medio	común	siempre existe	no (además de contar el número total de valores)	no	cuando hay valores atípicos, podría ser más representativa de un “promedio” que la media
Moda	valor más frecuente	usada algunas veces	podría no haber moda, una moda, o más de una moda	no	no	más apropiada para datos cualitativos (vea la sección 2.1)

La decisión de cómo tratar con los valores atípicos es uno de los temas más importantes en estadística. Algunas veces, como en nuestro ejemplo de básquetbol, un valor atípico es un valor legítimo que debe entenderse para interpretar la media y la mediana de manera apropiada. Otras veces, los valores atípicos indican errores en el conjunto de datos. Decidir cuándo los valores atípicos son importantes y cuándo sólo son errores puede ser muy difícil.

EJEMPLO 3 ¿Error?

Una entrenadora de atletismo quiere determinar un ritmo cardíaco apropiado para sus atletas durante sus sesiones de ejercicios. Ella selecciona a cinco de sus mejores corredores y les pide que utilicen monitores cardíacos durante un ejercicio. A la mitad del ejercicio lee los ritmos cardíacos de los cinco atletas: 130, 135, 140, 145, 325. En este caso, ¿cuál es la mejor medida del promedio, la media o la mediana? ¿Por qué?

Solución Cuatro de los cinco valores son muy cercanos y parece razonable para los ritmos cardíacos a mitad del ejercicio. El valor alto de 325 es un valor atípico. Lo más probable es que sea un error (quizás a causa de una falla en el monitor cardíaco), pues alguien con tan alto ritmo cardíaco debería sufrir un paro. Si la entrenadora utiliza la media como el promedio, incluiría este valor atípico, lo que significa que registraría cualquier error. Si ella utiliza la mediana como el promedio, tendría un valor más razonable, ya que la mediana no sería afectada por el valor atípico.

Confusión acerca de “promedio”

Los significados diferentes de *promedio* pueden conducir a confusión. Algunas veces esta confusión surge porque no decimos si el promedio es la media o la mediana, y otras veces porque no damos información suficiente de cómo se calculó el promedio. Los ejemplos siguientes ilustran unas situaciones.

EJEMPLO 4 Disputa salarial

Un periódico hizo una encuesta de salarios para trabajadores en compañías regionales de alta tecnología y reporta un promedio de \$22 por hora. De inmediato los trabajadores de una gran compañía solicitan un aumento en su salario, afirman que trabajan tan duro como los empleados de las otras compañías, pero su salario promedio es de sólo \$19. El administrador rechaza su petición, diciéndoles que están *muy bien pagados* ya que su salario promedio, de hecho, es de \$23. ¿Ambas partes pueden tener la razón? Explique.

Solución Ambas partes pueden estar en lo correcto, si utilizan definiciones diferentes de *promedio*. En su caso, los trabajadores podrían estar utilizando la mediana, mientras que el administrador utiliza la media. Por ejemplo, imagine que sólo hay cinco trabajadores y que sus salarios son \$19, \$19, \$19, \$19 y \$39. La mediana de estos cinco salarios es \$19 (como afirmaron los trabajadores), pero la media es \$23 (como afirmó el administrador).

EJEMPLO 5 ¿Cuál media?

Los 100 estudiantes de primer año en una pequeña universidad toman tres cursos en el programa básico de estudios. Dos cursos se imparten en grupos grandes, con los 100 estudiantes en un solo grupo. El tercer curso se imparte en 10 grupos de 10 estudiantes cada uno. Los estudiantes y los administradores argumentaron que los grupos son demasiado grandes. Los estudiantes aseguran que el tamaño medio de sus grupos en cursos básicos es 70. Los administradores aseguran que el tamaño medio del grupo es de sólo 25. ¿Ambas partes pueden estar correctas? Explique.

Solución Ambas partes están en lo correcto, pero hablan de medias diferentes. Los estudiantes calcularon el tamaño medio de los grupos en los que cada estudiante estaba inscrito de manera personal. Cada estudiante está tomando dos clases con 100 inscritos y una clase con 10 inscritos, por lo que el tamaño medio del grupo de cada estudiante es

$$\frac{\text{total de inscritos en las clases de un estudiante}}{\text{número de clases que el estudiante toma}} = \frac{100 + 100 + 10}{3} = 70$$

Los administradores calcularon la media de inscritos en todas las clases. Hay dos clases con 100 estudiantes y 10 clases con 10 estudiantes, haciendo un total de 300 estudiantes en 12 clases. La media de inscritos por clase es

$$\frac{\text{total de inscritos}}{\text{número de clases}} = \frac{300}{12} = 25$$

Las dos afirmaciones acerca de la media son correctas, pero muy diferentes porque los estudiantes y los administradores hablan de medias diferentes. Los estudiantes calcularon la media del *tamaño de grupo por estudiante*, mientras que los administradores calcularon el número medio de *estudiantes por clase*.

UN MOMENTO DE REFLEXIÓN

En el ejemplo 5, ¿los administradores podrían redistribuir las asignaciones a la facultad de modo que todas las clases tengan 25 estudiantes cada una? ¿Cómo? Analice las ventajas y las desventajas de tal cambio.

Media ponderada

Suponga que la calificación de su curso tiene como base cuatro cuestionarios y un examen final. Cada cuestionario cuenta 15% de su calificación final y el examen final cuenta 40%. Las calificaciones de sus cuestionarios son 75, 80, 84 y 88, y la calificación de su examen final es 96. ¿Cuál es su calificación global?

Puesto que el examen final cuenta más que los cuestionarios, una media simple de las cinco calificaciones no daría su calificación final. En lugar de eso debemos asignar un *peso* (que indique la importancia relativa) para cada calificación. En este caso asignamos pesos de 15 (para 15%) a cada uno de los cuestionarios y 40 (para 40%) al final. Luego, al sumar los productos de cada calificación con su peso y después dividirlos entre la suma de los pesos encontramos la **media ponderada**:

$$\begin{aligned} \text{media ponderada} &= \frac{(75 \times 15) + (80 \times 15) + (84 \times 15) + (88 \times 15) + (96 \times 40)}{15 + 15 + 15 + 15 + 40} \\ &= \frac{8745}{100} = 87.45 \end{aligned}$$

La media ponderada de 87.45 cuenta de manera apropiada los pesos diferentes de los cuestionarios y del examen. Siguiendo la regla de redondeo esta calificación queda en 87.5.

Las medias ponderadas son adecuadas siempre que los datos varíen de acuerdo con su grado de importancia. Siempre puede encontrar una media ponderada mediante la fórmula siguiente.

Definición

Una **media ponderada** toma en cuenta las variaciones en la importancia relativa de los valores. A cada valor se le asigna un peso y la media ponderada es

$$\text{media ponderada} = \frac{\text{suma de (cada valor} \times \text{su peso)}}{\text{suma de todos los pesos}}$$

A propósito...

Estadísticas deportivas que clasifican a los jugadores o equipos de acuerdo con su desempeño en muchas categorías diferentes por lo común son medias ponderadas. Ejemplos incluyen el promedio de carreras limpias y porcentaje de bateo en béisbol, el *rating* de los mariscales de campo en el fútbol americano y las clasificaciones computarizadas de equipos universitarios.

UN MOMENTO DE REFLEXIÓN

Puesto que los pesos son porcentajes, en el ejemplo de las calificaciones del curso podría considerar los pesos como 0.15 y 0.40 en lugar de 15 y 40. Calcule la media ponderada usando los pesos 0.15 y 0.40. ¿Obtiene la misma respuesta? ¿Por qué sí o por qué no?

EJEMPLO 6 CPG

Randall tiene 38 créditos con una calificación de A, 22 créditos con una calificación de B y 7 créditos con una calificación de C. ¿Cuál es su calificación promedio global (CPG)? Tome como base para la CPG valores de 4.0 para una A, 3.0 para una B y 2.0 puntos para una C.

Solución Las calificaciones de A, B y C representan valores de 4.0, 3.0 y 2.0, respectivamente. Los números de créditos son los pesos. Las A representan un valor de 4 con un peso de 38, las B representan un valor de 3 con un peso de 22 y las C representan un valor de 2 con un peso de 7. La media ponderada es

$$\text{media ponderada} = \frac{(4 \times 38) + (3 \times 22) + (2 \times 7)}{38 + 22 + 7} = \frac{232}{67} = 3.46$$

Siguiendo nuestra regla de redondeo, la CPG de Randall es 3.46 a 3.5.

EJEMPLO 7 Votos de accionistas

La votación en elecciones corporativas se pondera por la cantidad de acciones que posee cada votante. Suponga una compañía que tiene cinco accionistas quienes deciden si la compañía debe involucrarse en una nueva campaña publicitaria. Los votos (S = sí, N = no) son como sigue:

Accionista	Acciones en posesión	Voto
A	225	S
B	170	S
C	275	S
D	500	N
E	90	N

De acuerdo con los estatutos de la compañía, la medida necesita 60% de los votos para aprobarse. ¿Se aprueba?

Solución Podemos considerar un voto al “sí” como un valor de 1 y un voto al “no” como un valor 0. El número de acciones es el peso para el voto de cada accionista, por lo que el voto del accionista A representa un valor de 1 con un peso de 225, el voto del accionista B representa un valor de 1 con un peso de 170 y así sucesivamente. La media ponderada de los votos es

$$\begin{aligned} \text{media ponderada} &= \frac{(1 \times 225) + (1 \times 170) + (1 \times 275) + (0 \times 500) + (0 \times 90)}{225 + 170 + 275 + 500 + 90} \\ &= \frac{670}{1260} \approx 0.53 \end{aligned}$$

El voto ponderado es 53% (o 0.53) a favor, lo cual es inferior al 60% que se requiere, por lo que la medida no pasa.

Medias con notación de suma (sección opcional)

Muchas fórmulas estadísticas, incluyendo la fórmula para la media, pueden escribirse de manera compacta con una notación matemática denominada *notación de suma*. El símbolo Σ (letra griega *sigma* mayúscula) se denomina *signo de la suma* e indica que un conjunto de números debe sumarse. Utilizamos el símbolo x para representar *cada* valor en un conjunto de datos, por lo que escribimos la suma de todos los valores como

$$\text{suma de todos los valores} = \Sigma x$$

Por ejemplo, si una muestra consiste en 25 calificaciones de exámenes, Σx representa la suma de las 25 calificaciones. De manera análoga, si una muestra consiste en los ingresos de 10 000 familias, Σx representa el valor total de los 10 000 ingresos.

Utilizamos n para representar el número total de valores en la muestra. Así, la fórmula general para la media es

$$\bar{x} = \text{media muestral} = \frac{\text{suma de todos los valores}}{\text{número total de valores}} = \frac{\Sigma x}{n}$$

El símbolo $\bar{x}$ es el símbolo estándar para la media de una muestra. Cuando tratamos con la media de una población en lugar de una muestra, los estadísticos utilizan la letra griega μ (*mu*).

La notación de suma también hace más sencillo expresar una fórmula general para la media ponderada. Nuevamente utilizamos x para representar cada valor, y hacemos que w represente el peso de cada valor. La suma de los productos de cada valor por su peso correspondiente es $\Sigma(x \times w)$. La suma de los pesos es Σw . Por tanto, la fórmula para la media ponderada es

$$\text{media ponderada} = \frac{\Sigma(x \times w)}{\Sigma w}$$

Medias y medianas con datos que están en clases (sección opcional)

Las ideas de esta sección pueden extenderse a datos en clases con sólo suponer que el valor de en medio de la clase representa a todos los valores de la clase. Por ejemplo, considere la tabla siguiente de 50 valores separados en clases:

Clase	Frecuencia
0-6	10
7-13	10
14-20	10
21-27	20

El valor de la mitad de la primera clase es 3, por lo que suponemos que el valor 3 aparece 10 veces. Continuando de esta manera, tenemos para el total de 50 valores en la tabla

$$(3 \times 10) + (10 \times 10) + (17 \times 10) + (24 \times 20) = 780$$

Así, la media es $780/50 = 15.6$. Con 50 valores, la mediana está entre los valores 25 y 26. Estos valores caen dentro de la clase 14-20, por lo cual la llamamos **clase mediana** para los datos. La moda es la clase con la frecuencia más alta, que en este caso es la clase 21-27.

NOTA TÉCNICA

Las sumas con frecuencia se escriben con el uso de un *índice* que especifica cómo pasar a través de la suma. Por ejemplo, el símbolo x_i indica el i -ésimo dato en el conjunto; la letra i es el índice. Entonces escribimos la suma de todos los valores como

$$\sum_{i=1}^n x_i$$

Esta expresión la leemos como “la suma de los valores x_i , iniciando con $i = 1$ y continuando hasta $i = n$, donde n es el número total de valores en el conjunto”. Con esta notación, la media se escribe

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Sección 4.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Dato atípico.** ¿Qué es un dato atípico? ¿Los datos atípicos se definen de una manera exacta de modo que todos pueden ser identificados clara y objetivamente? Explique.
- Media y mediana.** Un grupo de estadística consiste de 24 estudiantes, de los cuales están desempleados o están empleados en trabajos de tiempo parcial con sueldos bajos. El grupo también incluye un profesor a quien se le paga un enorme sueldo. ¿Cuál describe mejor el ingreso de una persona típica en la clase de 25 personas (incluyendo al profesor): la media o la mediana? ¿Por qué?
- Tiempo medio de viaje al trabajo.** Una socióloga quiere determinar el tiempo medio de traslado al trabajo de todos los trabajadores residentes en Estados Unidos. Ella sabe que no es práctico encuestar a cada uno de los millones de trabajadores, así que realiza una búsqueda en Internet y encuentra la media del tiempo de traslado al trabajo para cada uno de los 50 estados. Suma los 50 tiempos de viaje al trabajo y los divide entre 50. ¿El resultado es una buena estimación del tiempo medio de viaje para todos los trabajadores? ¿Por qué sí o por qué no?
- Datos nominales.** En el capítulo 2 observó que los datos están en el nivel nominal de medida si sólo consisten en nombres o etiquetas. Un aficionado registra el número de la playera de cada jugador de los Patriots de Nueva Inglaterra en un juego del Súper Tazón. ¿Tiene sentido calcular la media de estos números? ¿Por qué sí o por qué no?

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Media.** Un conjunto de datos tiene medias de 65.2 y 72.3.
- Moda.** Un conjunto de datos tiene modas de 65.2 y 72.3.
- Media, mediana y moda.** Un investigador determina que un conjunto de datos tiene el mismo valor para la media, mediana y moda.
- Media ponderada.** Un investigador calcula el valor de la media para un conjunto de datos y luego construye una tabla de frecuencias y calcula la media a partir de la tabla de frecuencias. El investigador concluye que se cometió un error ya que se obtuvieron dos resultados diferentes.

Promedio apropiado. Los ejercicios 9 al 12 listan “promedios” que alguien podría querer conocer. En cada caso indique si la media o la mediana daría una mejor descripción del “promedio”. Explique su razonamiento.

- Ingreso.** El ingreso promedio de todos los adultos en una gran ciudad.
- Peso.** El peso promedio de las naranjas en una caja grande.
- Cambios de trabajo.** El número promedio de veces que las personas cambian de trabajo durante sus carreras.
- Equipaje perdido.** El número promedio de piezas de equipaje extraviado por vuelo para cada compañía aérea.

Conceptos y aplicaciones

Media, mediana y moda. Cada uno de los ejercicios 13 a 20 listan un conjunto de números. En cada caso determine la media, la mediana y la moda de los números listados.

- Percepción de tiempo.** Los tiempos reales (en segundos) registrados cuando estudiantes de estadística participaron en un experimento que probó su capacidad para determinar cuando había pasado un minuto (60 segundos):

53 52 75 62 68 58 49 49

- Temperaturas corporales.** Las temperaturas corporales (en grados Fahrenheit) de adultos normales seleccionados aleatoriamente:

98.6 98.6 98.0 98.0 99.0
98.4 98.4 98.4 98.4 98.6

- Alcohol en la sangre.** Concentraciones de alcohol de conductores involucrados en accidentes mortales y que luego recibieron sentencia de cárcel (con base en datos del Departamento de Justicia de Estados Unidos):

0.27 0.17 0.17 0.16 0.13 0.24
0.29 0.24 0.14 0.16 0.12 0.16

- Géiser el Viejo Fiel.** Intervalos de tiempo (en minutos) entre erupciones del Viejo Fiel en el Parque Nacional de Yellowstone:

98 92 95 87 96 90
65 92 95 93 98 94

- Moscas de la fruta.** Longitudes del tórax (en milímetros) de una muestra de moscas machos de la fruta:

0.72 0.90 0.84 0.68 0.84 0.90
0.92 0.84 0.64 0.84 0.76

- Edades de presidentes.** Edades de presidentes electos de Estados Unidos en el momento del inicio de su mandato:

57 61 57 57 58 57 61 54
68 51 49 64 50 48 65

- 19. Pesos de M&M.** Pesos (en gramos) de dulces M&M seleccionados aleatoriamente:

0.957 0.912 0.842 0.925 0.939 0.886
0.914 0.913 0.958 0.947 0.920

- 20. Monedas de 25 centavos.** Pesos (en gramos) de monedas de 25 centavos en circulación:

5.60 5.63 5.58 5.56 5.66 5.58 5.57 5.59
5.67 5.61 5.84 5.73 5.53 5.58 5.52 5.65
5.57 5.71 5.59 5.53 5.63 5.68

- 21. Estados en orden alfabético.** La tabla siguiente proporciona el área total, en millas cuadradas (tierra y agua) de los siete estados con nombres que inician con letras de la A a la C.

Estado	Área
Alabama	52 200
Alaska	615 200
Arizona	114 000
Arkansas	53 200
California	158 900
Colorado	104 100
Connecticut	5 500

- a. Determine el área media y el área mediana para estos estados.
- b. ¿Cuál estado es un dato atípico por la parte superior? Si elimina este estado, ¿cuál es la nueva área media y cuál el área mediana para este conjunto de datos?
- c. ¿Cuál estado es un dato atípico por la parte inferior? Si elimina este estado, ¿cuál es la nueva área media y cuál el área mediana para este conjunto de datos?
- 22. Dato atípico en cocas.** Las latas de Coca-Cola regular varían ligeramente en peso. A continuación están los pesos medidos de siete latas, en libras:
- 0.8161 0.8194 0.8165 0.8176
0.7901 0.8143 0.8126
- a. Determine la media y la mediana de estos pesos.
- b. ¿Cuál, si lo hay, de estos pesos usted consideraría que es un dato atípico? Explique.
- c. ¿Cuáles son los pesos medio y mediano, si se excluye el dato atípico?
- 23. Subir la calificación.** Suponga que usted tiene calificaciones de 70, 75, 80 y 70 en exámenes en una clase de matemáticas.
- a. ¿Cuál es la media de estas calificaciones?
- b. ¿Qué calificación necesitaría obtener en el siguiente examen para tener una media global de 75?
- c. Si la calificación máxima en un examen es 100, ¿es posible tener una media de 80 después del quinto examen? Explique.

- 24. Subir la calificación.** Suponga que usted tiene calificaciones de 60, 70, 65, 85 y 80 en exámenes en una clase de sociología.

- a. ¿Cuál es la media de estas calificaciones?
- b. ¿Qué calificación necesitaría obtener en el siguiente examen para tener una media global de 75?
- c. Si la calificación máxima en un examen es 100, ¿cuál es la máxima calificación media que podría obtener después del siguiente examen? Explique.

- 25. Nueva media.** Suponga que después de seis exámenes tiene una calificación media de 80 (de un posible de 100). Si obtiene una calificación de 90 en el examen siguiente, ¿cuál es su nueva media? ¿Cuál es la calificación media máxima que podría tener después del siguiente examen? ¿Cuál es la calificación media mínima que podría tener después del siguiente examen?

- 26. Nuevo promedio de bateo.** Suponga que después de 30 veces al bat una beisbolista tiene un promedio de bateo de 0.300. Si ella pega de hit en su siguiente turno al bat, ¿cuál sería su nuevo promedio de bateo? (El promedio de bateo es el número de hits dividido entre el número de veces al bat).

- 27. Comparación de promedios.** Suponga que los funcionarios escolares aseguran que el promedio de calificación en lectura para alumnos de cuarto grado en el distrito es 73 (de 100 posibles). Como director sabe que sus alumnos de cuarto grado tienen las calificaciones siguientes: 55, 60, 68, 70, 87, 88, 95. ¿Podría estar justificado que usted asegurase que las calificaciones de sus estudiantes están por arriba del promedio del distrito? Explique.

- 28. Comparación de promedios.** Suponga que la NBA (Asociación Nacional de Básquetbol) reporta que el promedio de altura de los jugadores de básquetbol es 6'8". Como entrenador sabe que los jugadores de su equipo titular tienen alturas de 6'5", 6'6", 6'6", 7'0" y 7'2". ¿Estaría justificado que asegurase que su equipo titular tiene una altura promedio por arriba de la NBA? Explique.

- 29. Duraznos promedio.** Una tienda tiene tres canastos de duraznos. Uno tiene 50 duraznos y pesa 18 libras, otro tiene 55 duraznos y pesa 22 libras, el tercero tiene 60 duraznos y pesa 24 libras. ¿Cuál es el peso medio de todos los duraznos? Explique.

- 30. Confusión de promedio.** Un instructor tiene un grupo con 25 estudiantes y ellos tuvieron una calificación media de 86% en el examen de mitad de semestre. El segundo grupo tiene 30 estudiantes y ellos tuvieron una calificación media de 84% en el mismo examen. ¿Se concluye que la calificación media para ambos grupos es 85%? Explique.

- 31. ¿Cuál media?** Los 300 estudiantes de una preparatoria toman los mismos cuatro cursos. Tres de los cursos son impartidos en 15 grupos de 20 estudiantes cada uno. El

cuarto curso es impartido en 3 grupos de 100 estudiantes cada uno. Explique cómo el director de la escuela podría asegurar que el tamaño medio del grupo es de 25 estudiantes, mientras que los padres disgustados podrían asegurar que el tamaño medio del grupo es de 40 estudiantes. ¿Cuál media considera que es una mejor descripción del tamaño del grupo?

32. Calificación final. Su calificación del curso tiene como base una que obtuvo a mitad de semestre y que vale 15% de su calificación final, un proyecto de clase vale 20% de su calificación final, un conjunto de tareas vale 40% de su calificación final y un examen final 25% de su calificación final. Su calificación de mitad de semestre es 75, la calificación de su proyecto es 90, su calificación de tareas es 85 y la calificación del examen final es 72. ¿Cuál es su calificación global final?

33. Promedio de bateo. Recuerde que para obtener un promedio de bateo en béisbol debe dividir el número total de *hits* entre el número total de veces al *bat* (sin tomar en cuenta las bases por bolas, errores y algunos otros casos especiales). Un jugador lleva 2 de 4 (2 *hits* en 4 veces al *bat*) en el primer juego, 0 de 3 en el segundo juego y 3 de 5 en el tercero. ¿Cuál es su promedio de bateo? ¿De qué manera este número es un “promedio”?

34. Promediar promedios. Suponga que un jugador tienen un promedio de bateo en muchos juegos de 0.400. En su siguiente juego lleva 2 de 4, que es un promedio de bateo de 0.500 para el juego. ¿Concluye que su nuevo promedio de bateo es $(0.400 + 0.500)/2 = 0.450$? Explique.

35. Porcentaje de huevos. Un inspector agrícola determina que 8% de los huevos examinados en una granja contienen salmonella y 12% de los huevos en otra granja contienen salmonella. ¿Se concluye que entre las dos granjas un total de 10% de los huevos examinados tienen salmonella? ¿Por qué sí o por qué no?

36. Promedio de imparables. Además del promedio de bateo, otra medida del desempeño en el béisbol es el promedio de imparables. Para determinar este promedio, un sencillo tiene un valor de 1 punto, un doble un valor de 2 puntos, un triple un valor de 3 puntos y un cuadrangular un valor de 4 puntos. El promedio de imparables de un jugador es el número total de puntos dividido entre el número total de veces al *bat* (sin tomar en cuenta las bases por bolas, errores y algunos otros casos especiales). Un jugador tiene tres sencillos en cinco veces al *bat* en el primer juego, un triple y un sencillo en cuatro veces al *bat* en el segundo juego y un doble y un cuadrangular en cinco veces al *bat* en el tercer juego.

- ¿Cuál es el promedio de bateo?
- ¿Cuál es el promedio de imparables?
- ¿Es posible que un promedio de imparables sea mayor que 1? Explique.

37. Votos de accionistas. Una compañía pequeña tiene cuatro accionistas. Uno tiene 400 acciones, el segundo tiene 300 acciones, el tercero tiene 200 acciones y el cuarto 100. En una votación sobre una nueva campaña de publicidad, el primer accionista vota SÍ y los otros tres votan NO. Explique cómo el resultado de la votación puede expresarse como un promedio ponderado. ¿Cuál es el resultado de la votación?

38. CPG. Un sistema común para calcular una calificación promedio global (CPG) asigna 4 puntos a una A, 3 puntos a una B, 2 puntos a una C y 1 punto a una D. ¿Cuál es la CPG de un estudiante que obtuvo una A en un curso con 5 créditos y una B, una C y una D, respectivamente, en cada uno de tres cursos con 3 créditos?

39. Fenotipos de chícharos. Un experimento buscó determinar si una deficiencia de dióxido de carbono en la tierra afecta los fenotipos de los chícharos. A continuación se listan los códigos de fenotipos donde 1 = amarillo claro, 2 = verde-suave, 3 = amarillo-arrugado y 4 = verde arrugado. ¿Pueden obtenerse las medidas de tendencia central para estos valores? ¿Los resultados tienen sentido?

2	1	1	1	1	1	1	4	1	2	2	1	2
3	3	2	3	1	3	1	3	1	3	2	2	

40. Centro de la población de Estados Unidos. Imagine tomar un enorme mapa de Estados Unidos y colocar pesos en él para representar dónde vive la población. El punto donde el mapa estaría equilibrado se denomina centro medio de la población. La figura 4.2 muestra cómo la ubicación del centro medio de la población se ha desplazado desde 1790 hasta 1990. Brevemente explique el patrón que se muestra en el mapa.



Figura 4.2 Centro medio de la población. Fuente: *Statistical Abstract of the United States*.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 4 en www.aw.com/bbt.

41. Datos de salario. Muchos sitios web ofrecen datos sobre salarios para diferentes profesiones. Determine el salario para una profesión que esté considerando. ¿Cuáles son los salarios medio y mediano para esta profesión? ¿Cómo se comparan estos salarios con los de otras profesiones que le interesen?

- 42. ¿El mensaje es la mediana?** Lea el artículo “The Median Isn’t the Message” de Stephen Jay Gould, que se encuentra en la web. Escriba unos cuantos párrafos en los que describa el mensaje que Gould trata de dar a entender. ¿Cómo este mensaje es importante para otros pacientes diagnosticados con cáncer?
- 43. Datos del ombligo.** Separe los datos recolectados en el ejercicio 23 de la sección 3.1. Luego construya una tabla de frecuencias y dibuje un histograma de la distribución. ¿Cuál es la media de la distribución? ¿Cuál es la mediana de la distribución? Una vieja teoría dice que, en promedio, la razón del ombligo de los humanos es la razón dorada: $(1 + \sqrt{5})/2$. ¿Esta teoría parece precisa en sus observaciones?

EN LAS NOTICIAS

- 44. Promedios diarios.** Cite tres ejemplos de promedios que usted trató en su vida (tal como promedio de calificaciones o promedio de bateo). En cada caso explique si el promedio es una media, una mediana o algún otro tipo de promedio. Describa brevemente cómo es útil para usted el promedio.
- 45. Promedios en las noticias.** Encuentre tres noticias recientes que hagan referencia a algún tipo de promedio. En cada caso explique si el promedio es una media, una mediana o algún otro tipo de promedio.

4.2 Formas de las distribuciones

Hasta este capítulo hemos estudiado cómo describir el centro de una distribución de datos cuantitativos con medidas tales como la media y la mediana. Ahora ponemos nuestra atención en la descripción de la *forma* global de una distribución. Podemos ver la forma completa de una distribución en una gráfica. Nuestro objetivo actual es describir la forma general en palabras. Aunque tales descripciones lleven menos información que la gráfica completa, son útiles. Nos centramos en tres características de una distribución: su número de modas, su simetría o sesgo y su variación.

Puesto que estamos interesados principalmente en las formas *generales* de las distribuciones, a menudo es más fácil examinar las gráficas que presentan curvas suaves que los conjuntos de datos originales. La figura 4.3 muestra tres ejemplos de esta idea, dos en que las distribuciones se muestran como histogramas y una en que la distribución se muestra como una gráfica de línea. Las curvas suaves aproximan estas distribuciones, pero no muestran todos sus detalles.

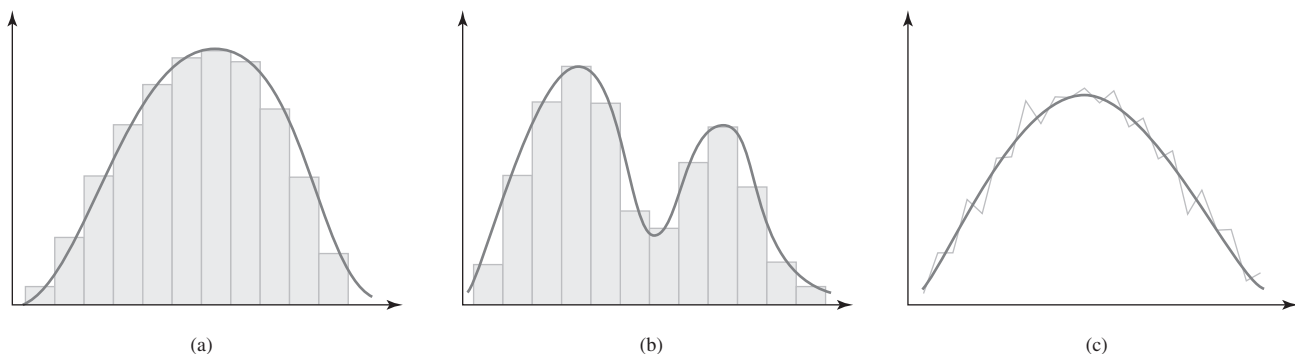


Figura 4.3 Las curvas suaves aproximan las formas de las distribuciones.

Número de modas

Una manera sencilla de describir la forma de una distribución es por su número de picos o modas. La figura 4.4a muestra una **distribución uniforme** que no tiene moda, ya que todos los valores tienen la misma frecuencia. La figura 4.4b muestra una distribución con un solo pico

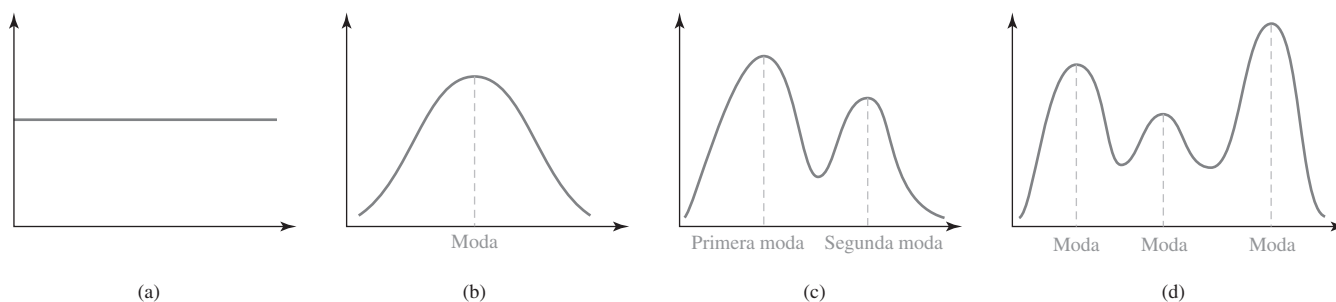


Figura 4.4 (a) Una distribución uniforme no tiene moda. (b) Una distribución con un solo pico tiene una moda. (c) Una distribución bimodal tiene dos modas. (d) Una distribución trimodal tiene tres modas.

como su moda. Se denomina distribución **con un solo pico** o **unimodal**. Por convención, cualquier pico en una distribución se considera una moda, incluso si no todos los picos tienen la misma altura. Por ejemplo, la distribución en la figura 4.4c tiene dos modas, aunque el segundo pico es más bajo que el primero; es una distribución *bimodal*. De manera análoga, la distribución en la figura 4.4d tiene tres modas, es una distribución *trimodal*.

EJEMPLO 1 Número de modas

¿Cuántas modas esperaría para cada una de las distribuciones siguientes? Haga un bosquejo para cada distribución, con los ejes claramente rotulados.

- Alturas de 1000 mujeres adultas elegidas aleatoriamente.
- Horas que adultos estadounidenses, seleccionados aleatoriamente, pasan viendo en televisión el fútbol americano en enero.
- Ventas semanales a lo largo del año en una tienda de ropa para niños.
- El número de personas con último dígito particular (0 a 9) en su número de seguridad social.

Solución La figura 4.5 muestra bosquejos de las distribuciones.

- La distribución de alturas de mujeres es con un solo pico, ya que muchas mujeres están en o cerca de la altura media, con cada vez menos mujeres en alturas mayores o en alturas menores que la media.
- La distribución de tiempo destinado a ver fútbol para 1000 adultos seleccionados aleatoriamente, quizá sea bimodal (dos modas). Una moda representa el tiempo medio que los hombres ven la televisión y la otra el tiempo medio de las mujeres.
- La distribución de las ventas semanales a lo largo del año en una tienda de ropa para niños quizá tenga varias modas. Por ejemplo, probablemente tendrá una moda en primavera para venta de ropa de verano, una moda al final del verano para ventas de regreso a la escuela y otra moda en invierno para ventas de fin de año.

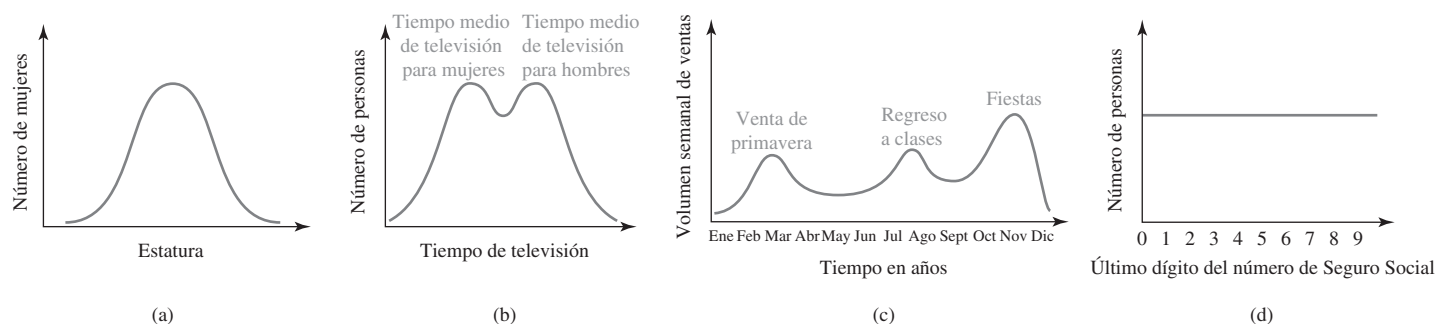


Figura 4.5 Bosquejos para el ejemplo 1.

- d. El último dígito de los números de seguridad social es aleatorio, por lo que el número de personas con cada último dígito diferente (0 a 9) debe ser casi la misma. Esto es, alrededor de 10% de todos los números de seguridad social terminan en 0, 10% terminan en 1, etcétera. Por tanto, es una distribución uniforme sin moda.

Simetría o sesgo

Una segunda forma sencilla para describir la forma de una distribución es en términos de su simetría o sesgo. Una distribución es **simétrica** si su mitad izquierda es una imagen en espejo de su mitad derecha. Las distribuciones de la figura 4.6 son simétricas. La distribución simétrica en la figura 4.6a, con un solo pico y una forma acampanada característica, es una *distribución normal*; es tan importante que dedicaremos el capítulo 5 a su estudio.

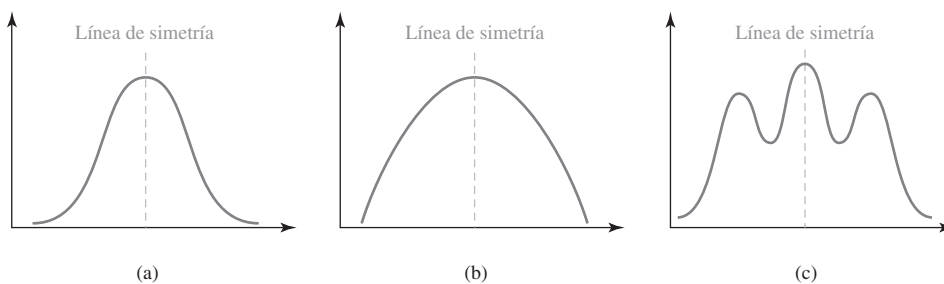


Figura 4.6 Todas estas distribuciones son simétricas, ya que sus mitades del lado izquierdo son imágenes de espejo de sus mitades del derecho. Observe que (a) y (b) tienen un solo pico (unimodal), mientras que (c) tiene tres picos (trimodal).

Una distribución que no es simétrica debe tener valores que tienden a desplegarse más en un lado que en el otro. En este caso decimos que la distribución es **sesgada**. La figura 4.7a muestra una distribución con valores más esparcidos en el lado izquierdo, lo que significa que algunos valores son atípicos en valores bajos. Decimos que tal distribución está **sesgada a la izquierda** (o sesgada *negativamente*). Es útil pensar que tal distribución tiene una cola que se jala hacia la izquierda. La figura 4.7b muestra una distribución con los valores más esparcidos a la derecha, haciéndola **sesgada a la derecha** (o sesgada *positivamente*). Parece como si la cola fuese jalada hacia la derecha.

La figura 4.7 también muestra cómo el sesgo afecta la posición relativa de la media, mediana y moda. Por definición, la moda está en el pico en una distribución con un solo pico.

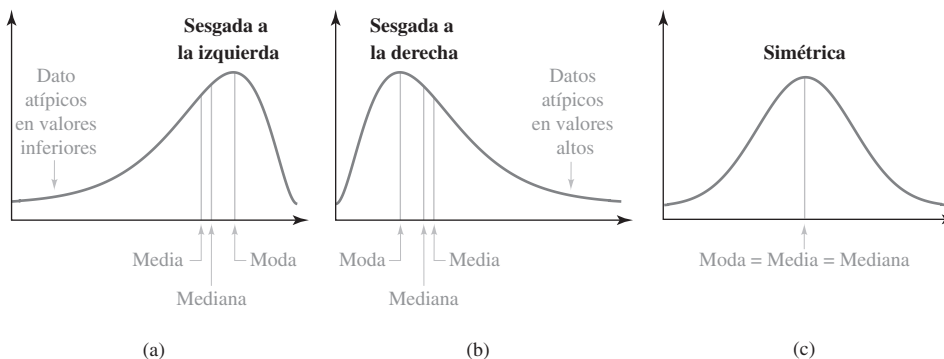


Figura 4.7 (a) Con sesgo a la izquierda (sesgada a la izquierda): la media y la mediana son menores que la moda. (b) Con sesgo a la derecha (sesgada a la derecha): la media y la mediana son mayores que la moda. (c) Distribución simétrica: la media, la mediana y la moda son iguales.

Una distribución sesgada a la izquierda jala la media y la mediana hacia la izquierda, esto es, a valores menores que la moda. Además, los valores atípicos en el extremo inferior del conjunto de datos tienden a hacer la media menor que la mediana (vea la tabla 4.3 en la página 149). De forma análoga, una distribución sesgada a la derecha jala a la media y a la mediana hacia la derecha, esto es, a valores mayores que la moda. En tales casos los valores atípicos en el extremo superior del conjunto de datos tienden a hacer la media mayor que la mediana. Cuando la distribución es simétrica y con un solo pico, tanto la media como la mediana son iguales a la moda.

NOTA TÉCNICA

Una distribución sesgada a la izquierda también se denomina *sesgada negativamente* y una distribución sesgada a la derecha también se denomina *sesgada positivamente*. Una distribución simétrica tiene *sesgo cero*.

Definiciones

Una distribución es **simétrica** si su mitad izquierda es una imagen de espejo de su mitad derecha.

Una distribución es **sesgada a la izquierda** si sus valores están más esparcidos en el lado izquierdo.

Una distribución es **sesgada a la derecha** si sus valores están más esparcidos en el lado derecho.

UN MOMENTO DE REFLEXIÓN

¿Cuál es una mejor medida del “promedio” (o del *centro* de la distribución) para una distribución sesgada: la mediana o la media? ¿Por qué?

EJEMPLO 2 Sesgo

Para cada una de las situaciones siguientes, indique si espera que la distribución sea simétrica, sesgada a la izquierda o sesgada a la derecha. Explique.

- Alturas de una muestra de 100 mujeres.
- Ingreso familiar en Estados Unidos.
- Velocidad de automóviles en una carretera donde está visible una patrulla que utiliza un radar para detectar conductores que van a alta velocidad.

Solución

- La distribución de alturas de mujeres es simétrica ya que aproximadamente igual número de mujeres son más bajas y más altas que la media, y alturas extremas son raras en cualquier lado de la media.
- La distribución de ingreso familiar es sesgada a la derecha. La mayoría son de clase media, por lo que la moda de esta distribución es un ingreso de la clase media (algo alrededor de \$50 000 en Estados Unidos). Pero unas cuantas familias jalan la media a un valor considerablemente más alto, estirando la distribución hacia el lado derecho (ingreso alto).
- Los conductores por lo regular reducen su velocidad cuando advierten la presencia de una patrulla que está buscando infractores del límite de velocidad. Pocos conductores excederán el límite de velocidad, pero algunos conductores tenderán a ir lento para asegurarse de ir abajo del límite de velocidad. Así, la distribución de velocidades será sesgada a la izquierda, con una moda cercana al límite de velocidad con algunos conductores muy abajo del límite de velocidad.

UN MOMENTO DE REFLEXIÓN

En español un sinónimo de *sesgado* es torcido y se utiliza para algo que está distorsionado o que se presenta de una manera injusta. En estadística, ¿cómo está relacionado el uso de *sesgo* con este significado?

Variación

Una tercera forma para describir una distribución es por medio de su **variación**, que es una medida de cuánto se esparcen los datos. Una distribución en la que la mayoría de los datos

A propósito...

La velocidad mata. En promedio, en Estados Unidos, cada 12 minutos alguien muere en un accidente automovilístico. Alrededor de una tercera parte de estos accidentes mortales involucran a un conductor a exceso de velocidad.

están muy cerca unos de otros tiene una variación baja. Como se muestra en la figura 4.8a, tal distribución tiene una forma muy puntiaguda. La variación es mayor cuando los datos están distribuidos más ampliamente alrededor del centro, lo cual hace más ancho el pico. La figura 4.8b muestra una distribución con una variación moderada y la figura 4.8c muestra una distribución con una variación alta. En la sección siguiente analizaremos métodos para describir la variación de manera cuantitativa.

Definición

La **variación** describe cuán ampliamente se distribuyen o están esparcidos los datos alrededor del centro de un conjunto de datos.

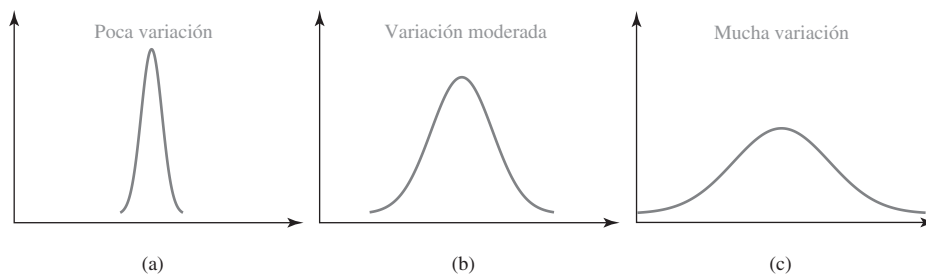


Figura 4.8 De izquierda a derecha estas tres distribuciones tienen variaciones en aumento.

EJEMPLO 3 Variación en tiempos del maratón

¿Cómo esperarías que difieran las variaciones entre los tiempos en el maratón olímpico y los tiempos en el maratón de la ciudad de Nueva York? Explique.

Solución El maratón olímpico sólo invita a corredores de élite, cuyos tiempos, con alta probabilidad están muy cerca de los récords mundiales. El maratón de la ciudad de Nueva York permite corredores de todas las capacidades, cuyos tiempos se distribuyen en un rango muy amplio (desde cerca del récord mundial, de un poco más de dos horas, a muchas horas). Por tanto, la variación entre los tiempos debe ser mayor en el maratón de la ciudad de Nueva York que en el maratón olímpico.

A propósito...

En 1997 los doctores le dijeron a Lance Armstrong, de 25 años de edad, que de acuerdo con los promedios, él tenía menos de 50-50 oportunidades de sobrevivir a su cáncer testicular, el cual ya se había esparcido a su cerebro. Él no sólo sobrevivió, fue a ganar la Vuelta a Francia (carrera en bicicleta) un total de siete veces consecutivas. Su historia muestra que, a nivel personal, la variación puede ser mucho más importante que cualquier promedio.



Sección 4.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Simetría.** ¿Qué queremos decir con una gráfica simétrica?
- 2. Distribución.** El voltaje suministrado a un hogar varía entre 123.0 volts y 125.0 volts, y todos los niveles de voltaje son igualmente probables. ¿Qué término describe mejor esta distribución: sesgada, bimodal, uniforme o unimodal?
- 3. Puntuación del IQ.** Considere las puntuaciones del IQ de estudiantes universitarios en un grupo de estadística comparado con el IQ de adultos seleccionados aleatoriamente. ¿Cuál de estos dos conjuntos de puntuaciones del IQ tienen menor variación? ¿Qué efecto tiene la menor variación sobre una gráfica de la distribución de esas puntuaciones del IQ?
- 4. Sesgo.** ¿Qué significa que una gráfica esté sesgada?
- 5. Simetría.** Puesto que un conjunto de datos tiene dos modas, no puede ser simétrico.
- 6. Simetría.** Una revisión del conjunto de datos revela que es simétrica, con una media de 75.0 y una mediana de 80.0.
- 7. Distribución uniforme.** Una revisión de un conjunto de datos muestra que su distribución es uniforme y unimodal.
- 8. Distribución.** Una revisión de un conjunto de datos revela que su distribución es sesgada a la izquierda y unimodal.

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

Conceptos y aplicaciones

- 9. El Viejo Fiel.** El histograma de la figura 4.9 muestra los tiempos entre erupciones del géiser Viejo Fiel del Parque Nacional de Yellowstone para una muestra de 300 erupciones (lo que significa 299 tiempos entre erupciones). Sobre el histograma dibuje una curva continua que capture sus características principales. Luego clasifique la distribución de acuerdo con su número de modas y su simetría o sesgo. Redacte un resumen del significado de sus resultados.

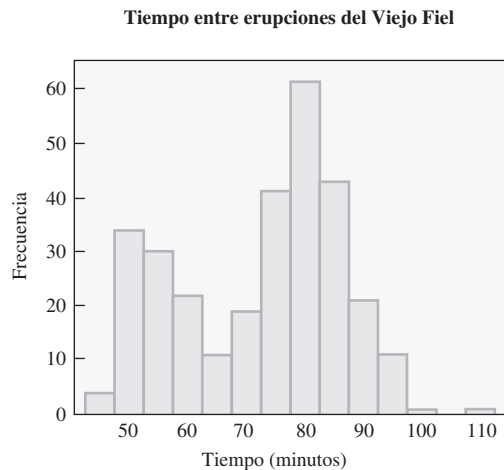


Figura 4.9 Fuente: Hand, et al. *Handbook of Small Data Sets*.

- 10. Fallas en chip.** El histograma de la figura 4.10 muestra los tiempos hasta la ocurrencia de una falla para una muestra de 108 circuitos de computadora. Sobre el histograma dibuje una curva continua que capture sus características principales. Luego clasifique la distribución de acuerdo con su número de modas y su simetría o sesgo. Redacte un resumen del significado de sus resultados.

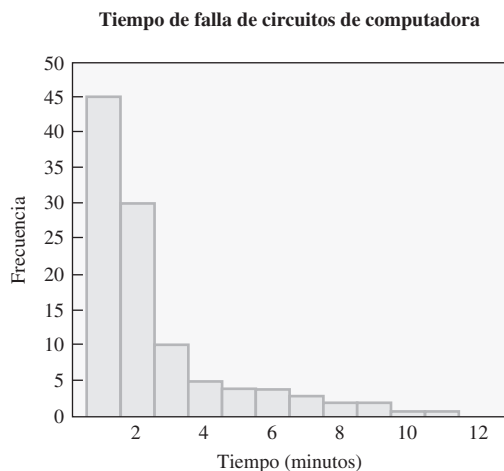


Figura 4.10 Fuente: Hand, et al. *Handbook of Small Data Sets*.

- 11. Pesos en el rugby.** El histograma de la figura 4.11 muestra los pesos de una muestra de 391 jugadores de rugby. Sobre el histograma dibuje una curva continua que capture sus características principales. Luego clasifique la distribución de acuerdo con su número de modas y su simetría o sesgo. Redacte un resumen del significado de sus resultados.

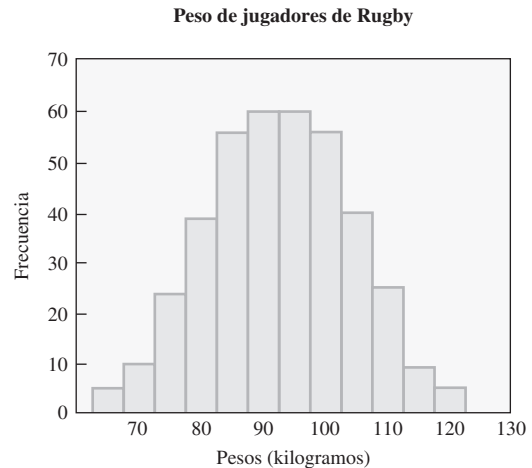


Figura 4.11 Fuente: Hand, et al. *Handbook of Small Data Sets*.

- 12. Pesos de monedas de un centavo.** El histograma de la figura 4.12 muestra los pesos (en gramos) de 72 monedas de un centavo. Sobre el histograma dibuje una curva continua que capture sus características principales. Luego clasifique la distribución de acuerdo con su número de modas y su simetría o sesgo. ¿Qué característica de la gráfica refleja el hecho que 35 de las monedas fueron acuñadas antes de 1983 y consisten en 97% de cobre y 3% de zinc, mientras que las otras 37 fueron acuñadas después de 1983 y tienen 3% de cobre y 97% de zinc.

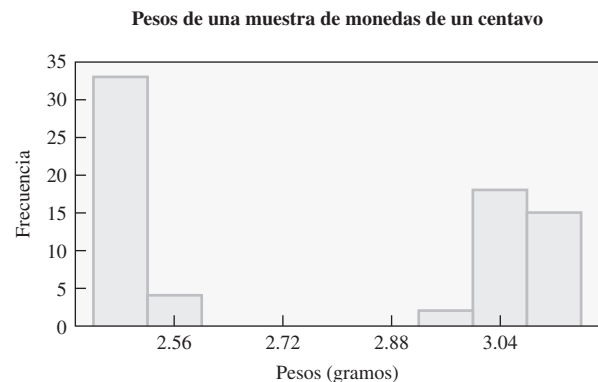


Figura 4.12 Fuente: Mediciones realizadas por Mario F. Triola.

- 13. Ingreso familiar.** Suponga que estudia el ingreso familiar en una muestra aleatoria de 300 familias. Sus resultados pueden resumirse como sigue.

- El ingreso familiar medio fue \$41 000.
 - El ingreso familiar mediano fue \$35 000.
 - Los ingresos más alto y más bajo fueron \$250 000 y \$2400, respectivamente.
- a. Haga un esbozo de la distribución de ingreso, con los ejes claramente rotulados. Describa la distribución como simétrica, sesgada a la derecha o la izquierda.
 - b. ¿Cuántas familias en la muestra ganan menos de \$35 000? Explique cómo lo sabe.
 - c. Con base en los datos, ¿puede determinar cuántas familias ganan más de \$41 000? ¿Por qué sí o por qué no?

14. Lluvia en Boston. La cantidad de lluvia diaria (en pulgadas) para Boston en un año reciente consistió de 365 valores con estas propiedades.

- La media de la cantidad de lluvia diaria es 0.083 pulgadas.
 - La mediana de la cantidad de lluvia diaria es 0 pulgadas.
 - La cantidad mínima de lluvia diaria es 0 pulgadas y la máxima es 1.48 pulgadas.
- a. ¿Cómo es posible que el mínimo de los 365 valores sea 0 pulgadas y la mediana también sea 0 pulgadas?
 - b. Describa la distribución como simétrica, sesgada a la izquierda o sesgada a la derecha.
 - c. ¿Puede determinar el número exacto de días que llovió? ¿Puede concluir algo acerca del número de días que llovió? Explique.

Descripción de distribuciones. Para cada distribución descrita en los ejercicios 15 al 30, responda las preguntas siguientes.

- a. ¿Cuántas modas esperaría para la distribución? Explique.
 - b. ¿Esperaría que la distribución sea simétrica, sesgada a la izquierda o sesgada a la derecha? Explique.
- 15. Ingresos.** Los ingresos anuales de todas aquellas personas que asisten a una clase de estadística, incluyendo al profesor.
- 16. Calificaciones de exámenes.** Las calificaciones de exámenes de 200 estudiantes que presentaron un buen examen con una calificación media de 75%.
- 17. Pesos de estudiantes.** Los pesos de 100 estudiantes de octavo grado.
- 18. Pesos de jugadores de fútbol americano.** Los pesos de 100 jugadores profesionales de fútbol americano.
- 19. Pesos de patinadores sobre hielo.** Los pesos de las personas que utilizaron una pista de hielo abierta únicamente para patinadores profesionales de figura artística en la mañana y para jugadores profesionales de hockey por la tarde y noche.

- 20. Pesos de vehículos.** Los pesos de automóviles en lote de automóviles nuevos de un distribuidor de automóviles, en los que aproximadamente la mitad del inventario consiste en automóviles compactos y la otra mitad a vehículos deportivos.
- 21. Demoras en las salidas.** Los retrasos en minutos de vuelos programados de un gran aeropuerto.
- 22. Velocidades.** Las velocidades de conductores cuando pasan por una zona escolar.
- 23. Edades de público.** Las edades de personas que visitan un museo de arte (sin incluir a grupos escolares).
- 24. Edades de público.** Las edades de personas que visitan un parque de diversiones temático.
- 25. Jugadores de tenis.** El número de jugadores en cada partido de dobles para mujeres en un torneo de tenis.
- 26. Ventas.** Las ventas mensuales de trajes de baño, durante un año, en una tienda de San Diego.
- 27. Ingresos.** Los ingresos de personas que en el Súper Tazón se sientan en lugares de lujo.
- 28. Ingresos.** Los ingresos de personas que ven el Súper Tazón por televisión.
- 29. Salarios.** Los salarios de jugadores de la Liga Mayor de Béisbol.
- 30. Promedios de bateo.** Los promedios de bateo de jugadores de la Liga Mayor de Béisbol.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 4 en www.aw.com/bbt.

- 31. Maratón de Nueva York.** El sitio web del maratón de la ciudad de Nueva York proporciona datos para los tiempos realizados en el más reciente maratón. Estudie los datos, construya un bosquejo de la distribución y describa la forma de la distribución.
- 32. Estadísticas de impuestos.** El sitio web IRS (Dirección General de Impuestos) proporciona estadísticas recolectadas de impuestos sobre el ingreso, devoluciones y mucho más. Seleccione un conjunto de estadísticas de este sitio web y estudie la distribución. Describa esta distribución y analice algo que vea y que sea relevante para la política nacional de impuestos.
- 33. Datos de seguridad social.** Una encuesta a una muestra de compañeros de estudio pide a cada uno indicar el último dígito de su número de seguridad social. También pide a cada participante indicar el quinto dígito. Dibuje una gráfica que muestre la distribución del último dígito y otra gráfica que muestre la distribución del quinto dígito. Compare las dos gráficas. ¿Qué diferencia notable se vuelve evidente?

EN LAS NOTICIAS

34. Distribución en las noticias. Encuentre tres ejemplos en las noticias de distribuciones que se muestren como histogramas o gráficas de líneas. En cada distribución dibuje una curva continua que capture sus características generales. Luego clasifique la distribución de acuerdo con su número de modas, simetría o sesgo y la variación.

35. Distribución trimodal. Proporcione un ejemplo de una distribución real que espera tenga *tres* modas.

Haga un bosquejo de la distribución, asegúrese de rotular los ejes en su bosquejo.

36. Distribución sesgada. Proporcione un ejemplo de una distribución real que esperaría sea sesgada a la derecha o sesgada a la izquierda. Construya un bosquejo de la distribución, asegúrese de rotular los ejes en su bosquejo.

4.3 Medidas de variación

En la sección 4.2 vimos la forma en que se describe la variación cualitativa. Aquí, pasamos a medidas cuantitativas de la variación.

¿Por qué interesa la variación?

*Nosotros los mortales cruzamos
el océano de este mundo.
Cada uno en su camarote
promedio de una vida.*

—Robert Browning

Imagine que observa a los clientes que esperan en la fila de cajeros en dos bancos diferentes. Los clientes del Gran Banco pueden formarse en cualquiera de tres filas diferentes, en la que cada una lleva a un cajero diferente. El Mejor Banco también tiene tres cajeros, pero todos los clientes esperan en una fila y son llamados hacia el siguiente cajero disponible. Los valores siguientes son tiempos de espera, en minutos, para 11 clientes en cada banco.

Gran Banco (tres filas): 4.1 5.2 5.6 6.2 6.7 7.2 7.7 7.7 8.5 9.3 11.0

Mejor Banco (una fila): 6.6 6.7 6.7 6.9 7.1 7.2 7.3 7.4 7.7 7.8 7.8

Probablemente encontrará clientes más descontentos en el Gran Banco que en el Mejor Banco, pero esto *no* es porque el promedio de espera sea mayor. De hecho, debe verificar que la media y la mediana de los tiempos de espera son 7.2 minutos en ambos bancos. La diferencia en la satisfacción de los clientes proviene de la *variación* en los dos bancos. Los tiempos de espera en el Gran Banco varían en un rango más amplio, por lo que unos cuantos clientes tienen largas esperas y es probable que se molesten. En contraste, la variación de los tiempos de espera en el Mejor Banco es menor, por lo que todos los clientes sienten que son tratados casi igualmente. La figura 4.13 muestra, con histogramas, la diferencia en las dos variaciones en las que los datos se clasifican al minuto más cercano.

A propósito...

La idea de esperar en línea (o *cola*) es importante no sólo para personas, sino también para datos, en particular para el flujo de información a través de internet. Con frecuencia grandes compañías emplean a estadísticos para ayudarles a asegurarse que la información fluye sin problemas y sin cuellos de botella por sus servidores y páginas web.

UN MOMENTO DE REFLEXIÓN

Explique *por qué* el Gran Banco con tres filas separadas debe tener una mayor variación en los tiempos de espera que el Mejor Banco. Luego considere varios lugares donde usted por lo común espera en filas, tal como un almacén, un banco, un punto de venta de boletos para un concierto o un restaurante de comida rápida. ¿Estos lugares utilizan una sola fila de clientes que van a varios cajeros o hay varias filas? Si un lugar utiliza varias filas, ¿considera que una sola fila sería mejor? Explique.

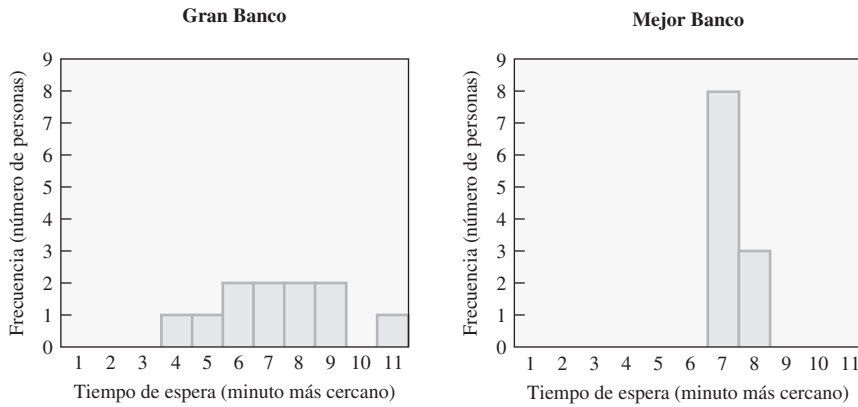


Figura 4.13 Histogramas para los tiempos de espera en el Gran Banco y en el Mejor Banco, mostrados con los datos agrupados en el minuto más cercano.

Rango

La manera más sencilla (pero no necesariamente la mejor) de describir la variación de un conjunto de datos es calcular su **rango**, definido como la diferencia entre los valores más pequeño (mínimo) y el más grande (máximo). Para el ejemplo de los dos bancos, los tiempos de espera para el Gran Banco varían desde 4.1 hasta 11.0 minutos, por lo que el rango es $11.0 - 4.1 = 6.9$ minutos. Los tiempos de espera para el Mejor Banco varían desde 6.6 a 7.8 minutos, por lo que el rango es $7.8 - 6.6 = 1.2$ minutos. El rango para el Gran Banco es mucho mayor, reflejado en su gran variación.

Definición

El **rango** de un conjunto de datos es la diferencia entre los valores máximo y mínimo:

$$\text{rango} = \text{valor más grande (máx)} - \text{valor más pequeño (mín)}$$

Aunque el rango es más fácil de calcular y puede ser útil, en ocasiones puede ser engañoso, como lo muestra el ejemplo siguiente.

EJEMPLO 1 Rango engañoso

Considere los dos conjuntos siguientes de calificaciones de exámenes rápidos para nueve estudiantes. ¿Cuál conjunto tiene el mayor rango? ¿También diría que este conjunto tiene la mayor variación?

Examen 1: 1 10 10 10 10 10 10 10 10
Examen 2: 2 3 4 5 6 7 8 9 10

Solución El rango para el examen 1 es $10 - 1 = 9$ puntos y el rango para el examen 2 es $10 - 2 = 8$ puntos. Así, el rango es mayor para el examen 1. Sin embargo, quitando la única calificación más baja (un dato atípico), el examen 1 no tiene variación ya que todos los demás estudiantes obtuvieron un 10. En contraste, ningún estudiante en el examen 2 obtuvo la misma calificación y las calificaciones se distribuyen en toda la lista de calificaciones posibles, por tanto, el examen 2 tiene mayor variación aunque el examen 1 tiene mayor rango.



© 1998 Scott Adams, Inc. Distribuido por United Feature Syndicate. Reimpreso con permiso. Todos los derechos reservados.

Cuartiles y el resumen de los cinco números

Una mejor manera de describir la variación es considerar algunos valores intermedios, además de los valores máximo y mínimo. Una manera incluye la búsqueda de **cuartiles**, o valores que dividen la distribución en cuartos. La lista siguiente repite los tiempos de espera en los dos bancos, con los cuartiles mostrados en negritas. Observe que el cuartil central, el cual divide al conjunto a la mitad, es simplemente la mediana.

	Cuartil inferior (Q_1)			Mediana (Q_2)			Cuartil superior (Q_3)			
			↓			↓			↓	
<i>Gran Banco:</i>	4.1	5.2	5.6	6.2	6.7	7.2	7.7	7.7	8.5	9.3 11.0
<i>Mejor Banco:</i>	6.6	6.7	6.7	6.9	7.1	7.2	7.3	7.4	7.7	7.8 7.8

NOTA TÉCNICA

Los estadísticos no coinciden acerca del procedimiento para calcular los cuartiles y procedimientos diferentes pueden resultar en valores un poco diferentes.

Definiciones

El **cuartil inferior** (o **primer cuartil** o Q_1) divide el cuarto inferior de los datos de los tres cuartos superiores. Es la mediana de los valores en la *mitad inferior* de un conjunto de datos. (Quitando el valor de en medio en el conjunto de datos, si el número de puntos es impar).

El **cuartil de en medio** (o **segundo cuartil** o Q_2) es la mediana global.

El **cuartil superior** (o **tercer cuartil** o Q_3) divide los tres cuartos de datos inferiores de un conjunto de datos del cuarto superior. Es la mediana de los valores en la *mitad superior* de un conjunto de datos. (Quitando el valor de en medio en el conjunto de datos, si el número de puntos es impar).

Una vez que conocemos los cuartiles, podemos describir una distribución con un **resumen de cinco números** que consiste del valor inferior, el cuartil inferior, la mediana, el cuartil superior y el valor máximo. Para los tiempos de espera en los dos bancos, el resumen de los cinco números es como sigue:

<i>Gran Banco:</i>		<i>Mejor Banco:</i>	
mínimo	= 4.1	mínimo	= 6.6
cuartil inferior	= 5.6	cuartil inferior	= 6.7
mediana	= 7.2	mediana	= 7.2
cuartil superior	= 8.5	cuartil superior	= 7.7
máximo	= 11.0	máximo	= 7.8

Resumen de los cinco números

El **resumen de los cinco números** para una distribución de datos consiste en los cinco números siguientes:

valor mínimo cuartil inferior mediana cuartil superior valor máximo

Podemos mostrar el resumen de los cinco números con una gráfica denominada **diagrama de cajas** o *boxplot* (algunas veces llamada *diagrama de cajas y bigotes*). Usando un número de referencia, encerramos los valores de los cuartiles inferior y superior en una caja. Luego dibujamos una línea que atraviese la caja en la mediana y agregamos dos “bigotes” que se extiendan de la caja a los valores máximo y mínimo. La figura 4.14 muestra el diagrama de cajas para los tiempos de espera de los bancos. Tanto la caja como los bigotes para el Gran Banco son más amplios que los del Mejor Banco, indicando que los tiempos de espera tienen mayor variación en el Gran Banco.

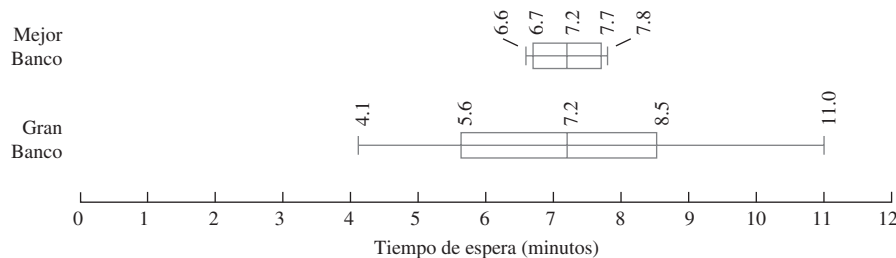


Figura 4.14 Los diagramas de cajas muestran que la variación de los tiempos de espera es mayor en el Gran Banco que en el Mejor Banco.

Cómo dibujar un diagrama de cajas

- Paso 1. Dibuje una recta numérica que se extienda a todos los valores en el conjunto de datos.
- Paso 2. Encierre los valores del cuartil inferior al cuartil superior en una caja. (El grosor de la caja no tiene relevancia).
- Paso 3. Dibuje una línea que cruce la caja en la mediana.
- Paso 4. Agregue “bigotes” que se extiendan a los valores mínimo y máximo.

NOTA TÉCNICA

Los diagramas de cajas mostrados en este libro se denominan *diagramas de cajas escuetos*. Algunos diagramas de cajas se dibujan con los datos atípicos marcados con un asterisco (*) y los bigotes se extienden sólo a los valores menores y mayores que no son valores atípicos; estos tipos de diagramas de cajas se denominan *diagramas de cajas modificados*.

EJEMPLO 2 Fumadores pasivos y activos

Una manera de estudiar la exposición al humo del cigarro es medir los niveles en la sangre de *suero de cotinina*, un producto metabólico de la nicotina que el cuerpo absorbe del humo del cigarro. La tabla 4.4 lista los niveles de suero de cotinina de muestras de 50 fumadores (“fumadores

Tabla 4.4 Niveles de suero de cotinina (nanogramos por milímetro de sangre) en muestras de 50 fumadores y 50 no fumadores expuestos a fumar pasivamente, con valores listados en orden ascendente					
Número de orden	Fumadores	No fumadores	Número de orden	Fumadores	No fumadores
1	0.08	0.03	26	34.21	0.82
2	0.14	0.07	27	36.73	0.97
3	0.27	0.08	28	37.73	1.12
4	0.44	0.08	29	39.48	1.23
5	0.51	0.09	30	48.58	1.37
6	1.78	0.09	31	51.21	1.40
7	2.55	0.10	32	56.74	1.67
8	3.03	0.11	33	58.69	1.98
9	3.44	0.12	34	72.37	2.33
10	4.98	0.12	35	104.54	2.42
11	6.87	0.14	36	114.49	2.66
12	11.12	0.17	37	145.43	2.87
13	12.58	0.20	38	187.34	3.13
14	13.73	0.23	39	226.82	3.54
15	14.42	0.27	40	267.83	3.76
16	18.22	0.28	41	328.46	4.58
17	19.28	0.30	42	388.74	5.31
18	20.16	0.33	43	405.28	6.20
19	23.67	0.37	44	415.38	7.14
20	25.00	0.38	45	417.82	7.25
21	25.39	0.44	46	539.62	10.23
22	29.41	0.49	47	592.79	10.83
23	30.71	0.51	48	688.36	17.11
24	32.54	0.51	49	692.51	37.44
25	32.56	0.68	50	983.41	61.33

Nota: La columna Número de orden se incluye para hacer más sencilla la lectura de la tabla.
Fuente: Encuesta Nacional de Examen de Salud y Nutrición, Instituto Nacional de Salud.

A propósito...

El tabaquismo pasivo parece ser particularmente perjudicial para los niños pequeños. Es evidente que las toxinas en el humo de los cigarrillos tienen un mayor efecto en cuerpos en desarrollo que en adultos completamente desarrollados. Un efecto similar se encuentra en la mayoría de las otras toxinas, por lo que es especialmente importante limitar la exposición de los niños a químicos tóxicos.

activos”) y 50 no fumadores quienes están expuestos al humo del cigarro en casa o en el trabajo (“fumadores pasivos”). Compare los dos conjuntos de datos (fumadores y no fumadores) con resúmenes de cinco números y los diagramas de cajas, y analice sus resultados.

Solución Los dos conjuntos de datos ya están en orden ascendente, lo que hace más sencillo construir el resumen de los cinco números. Cada uno tiene 50 puntos, por lo que la mediana se encuentra a la mitad entre los valores números 25 y 26. Para los fumadores, el 25o. y 26o. valores son 32.56 y 34.21, respectivamente, por lo que la mediana es

$$\frac{32.56 + 34.21}{2} = 33.385$$

Para los no fumadores, los valores 25o. y 26o. son 0.68 y 0.82, respectivamente, por lo que la mediana es

$$\frac{0.68 + 0.82}{2} = 0.75$$

El cuartil inferior es la mediana de la *mitad inferior* de los valores, que es el valor 13o. en cada conjunto. El cuartil superior es la mediana en la *mitad superior* de los valores, que es el valor 38o. en cada conjunto. Los resúmenes de los cinco números para los dos conjuntos de datos son como sigue:

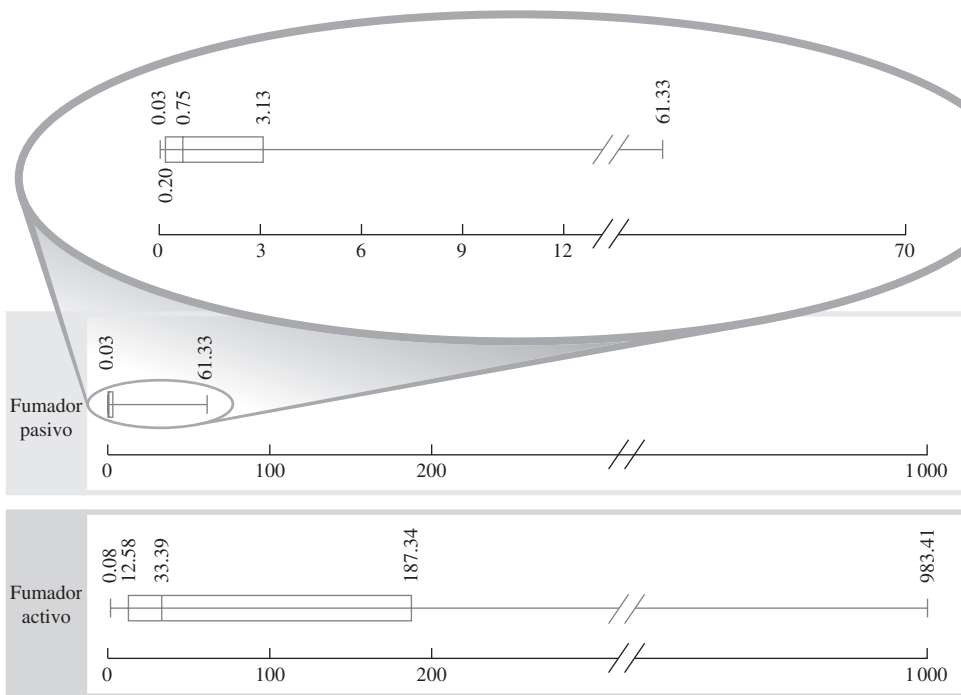


Figura 4.15 Diagramas de cajas para los datos en la tabla 4.4.

<i>Fumadores activos:</i>		<i>Fumadores pasivos:</i>	
mínimo	= 0.08 ng/ml	mínimo	= 0.03 ng/ml
cuartil inferior	= 12.58 ng/ml	cuartil inferior	= 0.20 ng/ml
mediana	= 33.385 ng/ml	mediana	= 0.75 ng/ml
cuartil superior	= 187.34 ng/ml	cuartil superior	= 3.13 ng/ml
máximo	= 983.41 ng/ml	máximo	= 61.33 ng/ml

La figura 4.15 muestra los diagramas de cajas para los dos conjuntos de datos. Los diagramas de cajas hacen más sencillo ver algunas de las características claves de los conjuntos de datos. Por ejemplo, es claro de inmediato que los fumadores activos tienen un nivel mediano más alto de suero de cotinina, así como una mayor variación en los niveles. Concluimos que los fumadores absorben considerablemente más nicotina que los no fumadores expuestos a tabaquismo pasivo. Sin embargo, los niveles en los fumadores pasivos son mucho más altos que los encontrados en personas que no tuvieron exposición a humo de cigarro (como lo demuestran otros datos, no mostrados aquí). En realidad, el fumador con el valor más alto para tabaquismo pasivo ha absorbido más nicotina que el fumador mediano. Concluimos que el tabaquismo pasivo expone a no fumadores a cantidades grandes de nicotina. Dados los peligros conocidos del humo de cigarro, estos resultados nos dan razones para estar preocupados por los posibles efectos a la salud del tabaquismo pasivo.

Percentiles

Los cuartiles dividen un conjunto de datos en cuatro segmentos. Es posible dividir los datos aún más. Por ejemplo en *quintiles* que dividen a un conjunto de datos en cinco segmentos y los *deciles* que dividen en 10 segmentos a un conjunto de datos. Es particularmente común dividir los conjuntos de datos en 100 segmentos usando **percentiles**. En términos informales,

NOTA TÉCNICA

Al igual que con los cuartiles, los estadísticos y varios programas de cómputo estadístico pueden utilizar procedimientos un poco diferentes para calcular los percentiles, lo que da como resultado valores un poco diferentes.

por ejemplo el percentil 35o. es un valor que separa al 35% inferior de un conjunto de datos del 65% superior. (Con mayor precisión, el percentil 35o. es mayor o igual al 35% de los valores y menor o igual que 65% de los datos).

Si un valor está entre dos percentiles, es común decir que el valor está *en* el percentil más bajo de ellos. Por ejemplo, si su calificación es mayor que 84.7% de todas las personas que hicieron un examen de ingreso a la universidad, decimos que su calificación está en el percentil 84o.

Definición

El ***n*-ésimo percentil** de un conjunto de datos divide al *n*% inferior de los valores del $(100 - n)\%$ superior. Un valor que está entre dos percentiles con frecuencia se dice que está *en* el percentil inferior. Puede aproximar el percentil de cualquier valor con la fórmula siguiente:

$$\text{percentil de un valor} = \frac{\text{número de valores menores que este valor}}{\text{número total de valores en el conjunto de datos}} \times 100$$

Existen diferentes procedimientos para determinar un valor que corresponda a un percentil dado, pero uno es determinar el valor *L-ésimo*, donde *L* es el producto del percentil (en forma decimal) y el tamaño de la muestra. Por ejemplo, con 50 valores muestrales el percentil 12 es alrededor de $0.12 \times 50 = 6$ (el sexto valor).

EJEMPLO 3 Percentiles de la exposición a humo

Responda las preguntas siguientes concernientes a los datos en la tabla 4.4.

- ¿Cuál es el percentil para el valor de 104.54 ng/ml para los fumadores?
- ¿Cuál es el percentil para el valor de 61.33 ng/ml para los no fumadores?
- ¿Cuál es el dato para el percentil 36 para los fumadores? ¿Y para los no fumadores?

Solución Los resultados siguientes son aproximaciones.

- El dato de 104.54 ng/ml para fumadores es el valor 35o. en el conjunto, lo que significa que 34 valores son menores a él. Así, su percentil es

$$\frac{\text{número de valores menores que 104.54 ng/ml}}{\text{número total de valores en el conjunto de datos}} \times 100 = \frac{34}{50} \times 100 = 68$$

En otras palabras, el valor 35o. marca el percentil 68o.

- El dato de 61.33 ng/ml para no fumadores es el 50o. y el más alto en el conjunto, lo que significa que 49 datos están abajo de él. Así, su percentil es

$$\frac{\text{número de valores menores que 61.33 ng/ml}}{\text{número total de valores en el conjunto}} \times 100 = \frac{49}{50} \times 100 = 98$$

En otras palabras, el valor más alto en este conjunto está en el percentil 98o.

- Ya que hay 50 datos en el conjunto, el percentil 36o. es alrededor de $0.36 \times 50 = 18$, el 18o. valor. Para los fumadores este valor es 20.16 ng/ml y para no fumadores es 0.33 ng/ml.

Desviación estándar

El resumen de los cinco números caracteriza bien la variación, pero los estadísticos con frecuencia prefieren describir la variación con un solo número. El número único que se usa con mayor frecuencia para describir la variación se denomina **desviación estándar**.

La desviación estándar es una medida de cuánto se esparcen los datos alrededor de la media de un conjunto de datos. Para calcular una desviación estándar primero encontramos la media y luego determinamos cuánto cada dato “se desvía” de la media. Considere nuestros datos de los bancos, en los que el tiempo medio de espera fue 7.2 minutos para ambos bancos. Un tiempo de espera de 8.2 minutos tiene una **desviación de la media** (o sólo **desviación**) de 8.2 minutos $-$ 7.2 minutos $=$ 1.0 minutos, lo que significa que es 1.0 minutos mayor que la media. Un tiempo de espera de 5.2 minutos tiene una desviación de la media de 5.2 minutos $-$ 7.2 minutos $=$ -2 minutos (2 minutos *negativos*), ya que es 2.0 minutos *menos* que la media.

La desviación estándar es una medida del promedio de todas las desviaciones de la media. Sin embargo, la media real de las desviaciones siempre es cero, ya que las desviaciones positivas equilibran de manera exacta las desviaciones negativas. Para evitar llegar siempre a un valor de cero, calculamos la desviación estándar encontrando primero una media de los *cuadrados* de las desviaciones (ya que los cuadrados siempre son positivos, o cero) y tomando la raíz cuadrada al final. (Por razones técnicas, dividimos la suma de los cuadrados entre el número total de valores *menos* 1).

NOTA TÉCNICA

En la determinación de la desviación estándar, cuando se trata con datos de una *muestra*, una parte del cálculo incluye la división de la suma de los cuadrados de las desviaciones entre el número total de valores *menos* 1. Cuando tratamos con *toda una población*, no restamos el 1. En este libro sólo utilizaremos la fórmula para una muestra.

Cómo calcular la desviación estándar

Para calcular la desviación estándar de cualquier conjunto de datos:

Paso 1. Calcule la media del conjunto de datos. Luego determine la desviación de la media de cada dato restando la media del dato. Es decir, para cada dato,

$$\text{desviación de la media} = \text{dato} - \text{media}$$

Paso 2. Determine los cuadrados (segunda potencia) de todas las desviaciones de la media.

Paso 3. Sume todos los cuadrados de las desviaciones de la media.

Paso 4. Divida esta suma entre el número total de datos *menos* 1.

Paso 5. La desviación estándar es la raíz cuadrada de este cociente.

En general, estos pasos producen la fórmula de la desviación estándar:

$$\text{desviación estándar} = \sqrt{\frac{\text{suma de (desviaciones de la media)}^2}{\text{número total de valores de datos} - 1}}$$

NOTA TÉCNICA

El resultado del paso 4 se denomina *varianza* de la distribución. En otras palabras, la desviación estándar es la raíz cuadrada de la varianza. Aunque la varianza se utiliza en muchos cálculos estadísticos avanzados, no la utilizaremos en este libro.

Observe que, como elevamos al cuadrado las desviaciones en el paso 3 y luego tomamos la raíz cuadrada en el paso 5, las unidades de la desviación estándar son las mismas que las unidades de los datos. Por ejemplo, si los datos tienen unidades de minutos, la desviación estándar también tiene unidades de minutos.

En principio, la fórmula de la desviación es relativamente sencilla, pero los cálculos reales se vuelven muy tediosos, salvo para conjuntos pequeños de datos. Como consecuencia, por lo regular se calcula con ayuda de una calculadora o una computadora (vea la sección Uso de la tecnología al final de este capítulo). Sin embargo, le será más fácil entender la fórmula de la desviación estándar si intenta algunos ejemplos en los que trabaje todos los cálculos a detalle.

EJEMPLO 4 Cálculo de la desviación estándar

Calcule las desviaciones estándar de los tiempos de espera del Gran Banco y del Mejor Banco.

Solución Seguimos los cinco pasos para calcular las desviaciones estándar. La tabla 4.5 muestra cómo organizar el trabajo en los primeros tres pasos. La primera columna para cada banco lista los tiempos de espera (en minutos), la segunda columna lista las desviaciones de la

Tabla 4.5 Cálculo de la desviación estándar

Gran Banco			Mejor banco		
Tiempo	Desviación (Tiempo – Media)	(Desviación) ²	Tiempo	Desviación (Tiempo – Media)	(Desviación) ²
4.1	4.1 – 7.2 = –3.1	(–3.1) ² = 9.61	6.6	6.6 – 7.2 = –0.6	(–0.6) ² = 0.36
5.2	5.2 – 7.2 = –2.0	(–2.0) ² = 4.00	6.7	6.7 – 7.2 = –0.5	(–0.5) ² = 0.25
5.6	5.6 – 7.2 = –1.6	(–1.6) ² = 2.56	6.7	6.7 – 7.2 = –0.5	(–0.5) ² = 0.25
6.2	6.2 – 7.2 = –1.0	(–1.0) ² = 1.00	6.9	6.9 – 7.2 = –0.3	(–0.3) ² = 0.09
6.7	6.7 – 7.2 = –0.5	(–0.5) ² = 0.25	7.1	7.1 – 7.2 = –0.1	(–0.1) ² = 0.01
7.2	7.2 – 7.2 = 0.0	(0.0) ² = 0.0	7.2	7.2 – 7.2 = 0.0	(0.0) ² = 0.0
7.7	7.7 – 7.2 = 0.5	(0.5) ² = 0.25	7.3	7.3 – 7.2 = 0.1	(0.1) ² = 0.01
7.7	7.7 – 7.2 = 0.5	(0.5) ² = 0.25	7.4	7.4 – 7.2 = 0.2	(0.2) ² = 0.04
8.5	8.5 – 7.2 = 1.3	(1.3) ² = 1.69	7.7	7.7 – 7.2 = 0.5	(0.5) ² = 0.25
9.3	9.3 – 7.2 = 2.1	(2.1) ² = 4.41	7.8	7.8 – 7.2 = 0.6	(0.6) ² = 0.36
11.0	11.0 – 7.2 = 3.8	(3.8) ² = 14.44	7.8	7.8 – 7.2 = 0.6	(0.6) ² = 0.36
		Suma = 38.46			Suma = 1.98

media (paso 1), y la tercera columna lista los cuadrados de las desviaciones (paso 2). Sumamos todas las desviaciones elevadas al cuadrado para encontrar la suma en la parte inferior de la tercera columna (paso 3). Observe que podemos calcular la desviación puesto que ya sabemos que el tiempo medio de espera es 7.2 minutos para los dos bancos. Para el paso 4 dividimos las sumas del paso 3 entre el número total de valores *menos* 1. Ya que existen 11 valores, dividimos entre 10:

$$\text{Gran Banco: } \frac{38.46}{10} = 3.846$$

$$\text{Mejor Banco: } \frac{1.98}{10} = 0.198$$

Por último, el paso 5 nos dice que las desviaciones estándar son las raíces cuadradas de los números en el paso 4:

$$\text{Gran Banco: } \text{desviación estándar} = \sqrt{3.846} = 1.96 \text{ minutos}$$

$$\text{Mejor Banco: } \text{desviación estándar} = \sqrt{0.198} = 0.44 \text{ minutos}$$

En otras palabras, aunque los bancos tienen el mismo tiempo medio de espera de 7.2 minutos, los tiempos de espera comunes tienden a estar 1.96 minutos alejados de esta media en el Gran banco, pero sólo 0.44 minutos alejados de esta media en el Mejor Banco. Nuevamente vemos que los tiempos de espera mostraron mayor variación en el Gran Banco, explicando por qué las filas en el Gran Banco disgustan a más clientes que a los del Mejor Banco.

UN MOMENTO DE REFLEXIÓN

Revise con cuidado las desviaciones individuales en la tabla 4.5 del ejemplo 4. ¿Las desviaciones estándar para los dos conjuntos de datos parecen “promedios” razonables de las desviaciones? Explique.

Interpretación de la desviación estándar

Una manera adecuada de desarrollar una comprensión profunda de la desviación estándar es considerar una aproximación denominado **rango de la regla empírica**, resumido en el recuadro siguiente.

El rango de la regla empírica

La desviación estándar se relaciona de *manera aproximada* con el rango de una distribución por medio del **rango de la regla empírica**:

$$\text{desviación estándar} \approx \frac{\text{rango}}{4}$$

Si conocemos el rango de una distribución (rango = máximo – mínimo), podemos utilizar esta regla para estimar la desviación estándar. La alternativa es que si conocemos la desviación estándar, podemos utilizar esta regla para estimar los valores máximo y mínimo como sigue:

$$\text{valor mínimo} \approx \text{media} - (2 \times \text{desviación estándar})$$

$$\text{valor máximo} \approx \text{media} + (2 \times \text{desviación estándar})$$

El rango de la regla empírica no funciona bien cuando los valores máximo o mínimo son datos atípicos.

El rango de la regla empírica funciona razonablemente bien para conjuntos de datos en que los valores están distribuidos más o menos uniformemente. No funciona bien cuando los valores máximo y mínimo son valores atípicos. Por tanto, debe utilizar su juicio para decidir si el rango de la regla empírica es aplicable en un caso particular, y en todos los casos recordar que el rango de la regla empírica proporciona aproximaciones, no resultados exactos.

EJEMPLO 5 Uso del rango de la regla empírica

Utilice el rango de la regla empírica para estimar las desviaciones estándar para los tiempos de espera en el Gran Banco y en el Mejor Banco. Compare las estimaciones con los valores reales encontrados en el ejemplo 4.

Solución Los tiempos de espera para el Gran Banco varían de 4.1 a 11.0 minutos, lo cual significa un rango de $11.0 - 4.1 = 6.9$ minutos. Los tiempos de espera para el Mejor Banco varían de 6.6 a 7.8 minutos para un rango de $7.8 - 6.6 = 1.2$ minutos. Así, el rango de la regla empírica proporciona las estimaciones siguientes para las desviaciones estándar:

$$\text{Gran Banco:} \quad \text{desviación estándar} \approx \frac{6.9}{4} = 1.7$$

$$\text{Mejor Banco:} \quad \text{desviación estándar} \approx \frac{1.2}{4} = 0.3$$

Las desviaciones estándar calculadas en el ejemplo 4 son 1.96 y 0.44, respectivamente. Para estos dos casos las estimaciones de los rangos mediante el rango de la regla empírica subestiman ligeramente las desviaciones estándar reales. Sin embargo, las estimaciones nos ponen en el camino correcto, mostrando que la regla es útil.

EJEMPLO 6 Estimación de un rango

Los estudios de millaje de gasolina de un BMW bajo condiciones de conducción variable muestran que proporciona 22 millas por galón con una desviación estándar de 3 millas por galón. Estime las cantidades de millajes de gasolina máxima y mínima que puede esperar bajo condiciones ordinarias de manejo.

Solución Con base en la regla empírica para el rango, los valores mínimo o máximo para el millaje de gasolina aproximadamente son

$$\text{valor mínimo} \approx \text{media} - (2 \times \text{desviación estándar}) = 22 - (2 \times 3) = 16$$

$$\text{valor máximo} \approx \text{media} + (2 \times \text{desviación estándar}) = 22 + (2 \times 3) = 28$$

El rango de millaje de gasolina para el automóvil tiene aproximadamente un mínimo de 16 millas por galón a un máximo de 28 millas por galón.

NOTA TÉCNICA

Otra manera de interpretar la desviación estándar utiliza una regla matemática denominada *Teorema de Chebyshev*, el cual establece que, para cualquier distribución de datos, al menos 75% de todos los datos caen dentro de dos desviaciones estándar de la media, y al menos 89% de todos los datos están a tres desviaciones estándar de la media. Aunque no utilizaremos este teorema en este libro, podría encontrarlo si toma otro curso de estadística.

A propósito...

Las tecnologías, tal como los convertidores catalíticos, han ayudado a reducir las cantidades de muchos contaminantes emitidos por los automóviles (por milla conducida), pero quemar menos gasolina es la única manera para reducir las emisiones de dióxido de carbono que causan calentamiento global. Ésta es una de las razones principales por lo que los fabricantes de automóviles están desarrollando vehículos híbridos de alto millaje y vehículos de cero emisiones de funcionamiento continuo con electricidad o celdas de combustible.



Desviación estándar con notación de suma (sección opcional)

La notación de suma introducida anteriormente hace más sencillo escribir la fórmula de la desviación estándar en una forma compacta. Recuerde que x representa los valores individuales en un conjunto de datos y $\bar{x}$ representa la media del conjunto de datos. Por tanto, podemos escribir la desviación de la media para cualquier dato como

$$\text{desviación} = \text{dato} - \text{media} = x - \bar{x}$$

Ahora podemos escribir la suma de los cuadrados de todas las desviaciones como

$$\text{suma de todos los cuadrados de las desviaciones} = \sum (x - \bar{x})^2$$

Los pasos restantes en el cálculo de la desviación estándar son dividir esta suma entre $n - 1$ y luego tomar la raíz cuadrada. Debe confirmar que la fórmula siguiente resume los cinco pasos en el recuadro anterior:

$$s = \text{desviación estándar} = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

NOTA TÉCNICA

La fórmula para la *varianza* es

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

El símbolo estándar para la varianza, s^2 , refleja que es el cuadrado de la desviación estándar.

El símbolo s es el símbolo convencional para la desviación estándar de una muestra. Para la desviación estándar de una población, los estadísticos utilizan la letra griega σ (*sigma*) y el término $n - 1$ en la fórmula es reemplazada por n . En consecuencia, obtendrá resultados ligeramente diferentes para la desviación estándar dependiendo si supone que los datos representan una muestra o una población.

Sección 4.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Variación.** ¿Por qué la desviación estándar es considerada una medida de la variación? Con sus palabras describa la característica de un conjunto de datos que se mide con la desviación estándar.
- Comparación de la variación.** ¿Cuál considera que tiene más variación: las puntuaciones del CI de 30 estudiantes en una clase de física o las puntuaciones del CI de 30 personas en un cine? ¿Por qué?
- ¿Enunciado correcto?** En el libro *Cómo mentir con diagramas*, el autor escribe que “La desviación estándar por lo regular se muestra como más o menos la diferencia entre el máximo y la media, y el mínimo y la media. Por ejemplo, si la media es 1, el máximo es 3 y el mínimo es -1 , la desviación estándar es ± 2 ”. ¿El enunciado es correcto? ¿Por qué sí o por qué no?
- Cuartiles.** Para un conjunto grande de datos, el cuartil inferior se determinó y es 93.2. ¿Qué queremos decir cuando decimos que 93.2 es el cuartil inferior?

¿Tiene sentido? Para los ejercicios 5 al 8, decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Calificaciones del SAT.** Una calificación del SAT estuvo en el valor mediano y fue el 60o. percentil.
- Ingreso familiar.** Un reporte en *USA Today* indica que para un año reciente el ingreso mediano anual de una familia es \$43 200 y la media es \$70 700.
- Ingresos.** La desviación estándar de los ingresos anuales de 50 maestros de estadística de tiempo completo es menor que la desviación estándar de los ingresos anuales de 50 médicos.
- Tamaños de muestras.** La desviación estándar de los ingresos anuales de 50 maestros de estadística de tiempo completo elegidos aleatoriamente es mucho menor que la desviación estándar de los ingresos anuales de otros 200 maestros de estadística de tiempo completo elegidos aleatoriamente.

Conceptos y aplicaciones

Rango y desviación estándar. Cada uno de los ejercicios 9 a 16 listan un conjunto de números. En cada caso determine el rango y la desviación estándar. (Los mismos conjuntos de números fueron utilizados en los ejercicios 13 a 20 en la sección 4.1).

- Percepción del tiempo.** Los tiempos reales (en segundos) registrados cuando estudiantes de estadística participaron

en un experimento para probar su habilidad para determinar cuándo había pasado un minuto (60 segundos):

53 52 75 62 68 58 49 49

- 10. Temperaturas corporales.** Las temperaturas corporales (en grados Fahrenheit) de adultos normales y saludables elegidos aleatoriamente:

98.6 98.6 98.0 98.0 99.0
98.4 98.4 98.4 98.4 98.6

- 11. Alcohol en la sangre.** Concentraciones de alcohol en la sangre de conductores involucrados en accidentes mortales automovilísticos y luego sentenciados a prisión (con base en datos del Departamento de Justicia de Estados Unidos):

0.27 0.17 0.17 0.16 0.13 0.24
0.29 0.24 0.14 0.16 0.12 0.16

- 12. Géiser el Viejo Fiel.** Los intervalos de tiempo (en minutos) entre erupciones del géiser el Viejo Fiel en el Parque Nacional de Yellowstone:

98 92 95 87 96 90
65 92 95 93 98 94

- 13. Moscas de la fruta.** Longitudes del tórax (en milímetros) de una muestra de moscas, machos, de la fruta:

0.72 0.90 0.84 0.68 0.84 0.90
0.92 0.84 0.64 0.84 0.76

- 14. Edades de presidentes.** Edades de presidentes electos de Estados Unidos en el momento del inicio de su mandato:

57 61 57 57 58 57 61 54
68 51 49 64 50 48 65

- 15. Pesos de M&M.** Pesos (en gramos) de dulces M&M seleccionados aleatoriamente:

0.957 0.912 0.842 0.925 0.939 0.886
0.914 0.913 0.958 0.947 0.920

- 16. Monedas de 25 centavos.** Pesos (en gramos) de monedas de 25 centavos en circulación:

5.60 5.63 5.58 5.56 5.66 5.58 5.57 5.59
5.67 5.61 5.84 5.73 5.53 5.58 5.52 5.65
5.57 5.71 5.59 5.53 5.63 5.68

Comparación de la variación. En los ejercicios 17 a 20, determine el rango y la desviación estándar para cada una de las dos muestras y luego compare los dos conjuntos de resultados.

- 17. Lluven gatos.** En ocasiones las estadísticas son utilizadas para comparar o identificar autores de trabajos diferentes. Las longitudes de las primeras 20 palabras en el prólogo de Tennessee Williams en *Cat on a Hot Tin Roof* se listan junto con las longitudes de las primeras 20 palabras en *The Cat in the Hat* del Dr. Seuss. ¿Parece haber una diferencia en la variación?

Cat on a Hot Tin Roof:

2 6 2 2 1 4 4 2 4 2
3 8 4 2 2 7 7 2 3 11

The Cat in the Hat:

3 3 3 3 5 2 3 3 3 2
4 2 2 3 2 3 5 3 4 4

- 18. Edades de polizones.** El *Queen May* navegaba entre Inglaterra y Estados Unidos y en ocasiones se encontraban polizones a bordo. Las edades (en años) de ellos cuando iba hacia el este y cuando iba al oeste se dan a continuación (datos de Cunard Steamship, Co. Ltd.). Compare la variación en los dos conjuntos de datos.

Hacia el este:

24 24 34 15 19 22 18 20 20 17

Hacia el oeste:

41 24 32 26 39 45 24 21 22 21

- 19. Precisión en el pronóstico del clima.** En un análisis de la precisión de los pronósticos del clima, las temperaturas máximas reales se compararon con las temperaturas máximas pronosticadas un día antes y las temperaturas máximas pronosticadas cinco días antes. A continuación se listan los errores entre las temperaturas pronosticadas y las temperaturas máximas reales para días consecutivos en el condado de Dutchess, Nueva York. ¿Las desviaciones estándar sugieren que las temperaturas pronosticadas con un día de anticipación son más precisas que las pronosticadas con cinco días de anticipación, qué podría esperar?

(máxima real) – (máxima pronosticada un día antes)

2 2 0 0 -3 -2 1
-2 8 1 0 -1 0 1

(máxima real) – (máxima pronosticada cinco días antes)

0 -3 2 5 -6 -9 4
-1 6 -2 -2 -1 6 -4

- 20. Efecto de tratamiento.** Investigadores en la Universidad Estatal de Pennsylvania realizaron experimentos con álamos. A continuación se listan los pesos (en kilogramos) de los álamos a los que no se les dio el tratamiento y los álamos tratados con fertilizante e irrigación. ¿Parece haber una diferencia entre las dos desviaciones estándar?

Sin tratamiento:

0.15 0.02 0.16 0.37 0.22

Fertilizante e irrigación:

2.03 0.27 0.92 1.07 2.38

- 21. Cálculo de percentiles.** Uno de los autores con mucho tiempo de ocio pesó cada uno de los 465 dulces M&M de una bolsa.

- a. Uno de los M&M pesó 0.776 gramos y pesó más que 25 de los otros M&M. ¿Cuál es el percentil de este valor particular?
- b. Uno de los M&M pesó 0.876 gramos y fue más pesado que 322 de los otros M&M. ¿Cuál es el percentil de este valor particular?
- c. Uno de los M&M pesó 0.856 gramos y fue más pesado que 224 de los otros M&M. ¿Cuál es el percentil de este valor particular?

22. Cálculo de percentiles. Un conjunto de datos consiste en las edades de 76 mujeres en el momento que ganaron el Oscar en la categoría de mejor actriz.

- a. Una de las actrices tenía 34 años de edad y era mayor que 38 de las otras actrices en el momento que ganaron el Oscar. ¿Cuál es el percentil de la edad de 34?
- b. Una de las actrices tenía 29 años de edad y ella era mayor que 20 de las otras actrices en el momento que ganaron el Oscar. ¿Cuál es el percentil de la edad de 29?
- c. Una de las actrices tenía 60 años de edad y era mayor que 71 de las otras actrices en el momento que ganaron el Oscar. ¿Cuál es el percentil de la edad de 60?

23. Comprensión de la desviación estándar. Los cuatro conjuntos de datos siguientes de 7 números tienen una media de 9.

{9, 9, 9, 9, 9, 9, 9}, {8, 8, 9, 9, 9, 10, 10}
 {8, 8, 8, 9, 10, 10, 10}, {6, 6, 6, 9, 12, 12, 12}

- a. Construya un histograma para cada conjunto.
- b. Proporcione el resumen de los cinco números y dibuje un diagrama de cajas para cada conjunto.
- c. Compare las desviaciones estándar para cada conjunto.
- d. Con base en sus resultados, explique brevemente cómo la desviación estándar proporciona un resumen útil, con un solo número, de la variación en estos conjuntos de datos.

24. Comprensión de la desviación estándar. Los cuatro conjuntos de datos siguientes con 7 números tienen una media de 6.

{6, 6, 6, 6, 6, 6, 6}, {5, 5, 6, 6, 6, 7, 7}
 {5, 5, 5, 6, 7, 7, 7}, {3, 3, 3, 6, 9, 9, 9}

- a. Construya un histograma para cada conjunto.
- b. Proporcione el resumen de los cinco números y dibuje un diagrama de cajas para cada conjunto.
- c. Compare las desviaciones estándar para cada conjunto.
- d. Con base en sus resultados, explique brevemente cómo la desviación estándar proporciona un resumen útil, con un solo número, de la variación en estos conjuntos de datos.

Comparación de variaciones. Para cada uno de los ejercicios del 25 al 28, haga lo siguiente:

- a. Determine la media, mediana y el rango para cada uno de los conjuntos de los dos conjuntos.
 - b. Proporcione el resumen de los cinco números y dibuje un diagrama de cajas para cada uno de los dos conjuntos de datos.
 - c. Determine la desviación estándar para cada uno de los dos conjuntos de datos.
 - d. Aplique el rango de la regla empírica para estimar la desviación estándar de cada uno de los dos conjuntos de datos. ¿Qué tan bien funciona la regla en cada caso? Analice brevemente por qué si o por qué no funciona bien.
 - e. Con base en todos sus resultados, compare y analice los dos conjuntos de datos en términos de sus medidas centrales y su variación.
- 25.** Los conjuntos siguientes dan las edades, en años, de una muestra de automóviles en un estacionamiento para profesores y en un estacionamiento para estudiantes en el Colegio de Portland.

Profesores:

2 3 1 0 1 2 4 3 3 2 1

Estudiantes:

5 6 8 2 7 10 1 4 6 10 9

26. Los datos siguientes proporcionan las velocidades, en millas por hora, de los primeros nueve automóviles que pasan por una zona escolar y los primeros nueve automóviles que pasan por un cruce en el centro de la ciudad.

Escuela:

20 18 23 21 19 18 17 24 25

Centro:

29 31 35 24 31 26 36 31 28

27. Los datos siguientes muestran las edades de los primeros siete presidentes de Estados Unidos (desde Washington hasta Jackson) y de siete presidentes recientes de Estados Unidos (desde Nixon hasta G. W. Bush) en el momento de la toma del poder.

Primeros siete:

57 61 57 57 58 57 61

Siete recientes:

56 61 52 69 64 46 54

28. Los conjuntos de datos siguientes proporcionan las longitudes aproximadas de las nueve sinfonías de Beethoven y las nueve sinfonías de Mahler (en minutos),

Beethoven:

28 36 50 33 30 40 38 26 68

Mahler:

52 85 94 50 72 72 80 90 80

- 29. Entrega de pizzas.** Después de registrar los tiempos de entrega para dos diferentes tiendas de pizzas, usted concluye que una tienda tiene un tiempo medio de entrega de 45 minutos con una desviación estándar de 3 minutos, y la otra tienda tiene un tiempo medio de entrega de 42 minutos con una desviación estándar de 20 minutos. Interprete estas cifras. Si a usted le gustan por igual las pizzas de ambas tiendas, ¿a cuál ordenaría sus pizzas? ¿Por qué?
- 30. Quejas al director.** Usted administra una pequeña heladería en la que sus empleados sirven el helado a mano. Cada noche totaliza sus ventas y el volumen total de helado vendido. Encuentra que en las noches en las que un empleado, de nombre Sam, está trabajando, el precio medio del helado vendido es \$1.75 por porción, con una desviación estándar de \$0.05. En las noches que un empleado de nombre Ken está trabajando el precio medio del helado vendido es \$1.70 con una desviación estándar de \$0.35. ¿Cuál empleado es más probable que genere quejas de porciones “demasiado pequeñas”?
- 31. Desviación estándar de un portafolio.** El libro *Finanzas*, de Zvi Bodie, Alex Kane y Alan Marcus, asegura que los rendimientos para portafolios de inversión con una sola acción tienen una desviación estándar de 0.55, mientras que los rendimientos con 32 acciones tienen una desviación estándar de 0.325. Explique cómo la desviación estándar mide el riesgo en estos tipos de portafolios.
- 32. Desviación estándar de bateo.** Durante los últimos 100 años, la media del promedio de bateo en las ligas mayores ha permanecido constante alrededor de 0.260. Sin embargo, la desviación estándar de los promedios de bateo ha disminuido de casi 0.049 en el año 1870 a 0.031 en la actualidad. ¿Esto qué nos dice acerca de los promedios de bateo de los jugadores? Con base en estos hechos, ¿esperaría que los promedios de bateo por arriba de 0.350 sean más o menos comunes que en el pasado? Explique.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 4 en www.aw.com/bbt.

- 33. Tabaquismo pasivo.** En los sitios web de la Asociación Americana de Pulmones y la Agencia de Protección al

Medio Ambiente, encuentra datos estadísticos concernientes a los efectos en la salud del tabaquismo de segunda mano (pasivo). Escriba un resumen breve de sus hallazgos y sus opiniones acerca de si este tema de salud debe abordarse, y cómo, por parte del gobierno.

- 34. Los niños y los medios.** Un estudio reciente de la fundación de la familia Kaiser investigó el papel de los medios (por ejemplo, televisión, libros, computadoras) en las vidas de los niños. El reporte, que se encuentra en el sitio web de la fundación de la familia Kaiser, proporciona muchas distribuciones de datos respecto a, por ejemplo, cuánto tiempo destinan los niños diariamente a cada medio. Estudie al menos tres de las distribuciones y el reporte que encuentre particularmente interesante. Resuma cada distribución y analice sus opiniones de las consecuencias sociales de los hallazgos.
- 35. Medición de la variación.** El rango y la desviación estándar utilizan diferentes enfoques para medir la variación en un conjunto de datos. Construya dos conjuntos de datos configurados de modo que el rango del primer conjunto sea *mayor que* el rango del segundo conjunto (lo que sugiere que el primer conjunto tiene más variación), pero que la desviación estándar del primer conjunto sea *menor que* la desviación estándar del segundo conjunto (lo que sugiere que el primer conjunto tiene menos variación).

EN LAS NOTICIAS

- 36. Rangos en las noticias.** Encuentre dos ejemplos de distribuciones de datos en noticias recientes; pueden ser dadas como tablas o bien como gráficas. En cada caso indique el rango de la distribución y explique su significado en el contexto de la noticia. Estime la desviación estándar aplicando el rango de la regla empírica.
- 37. Resumen de un conjunto de datos en una noticia.** Encuentre un ejemplo de una distribución de datos dada en forma de una tabla en una noticia reciente. Construya un resumen de los cinco números y un diagrama de cajas para la distribución.

4.4 Paradojas estadísticas

El polígrafo (“detector de mentiras”) se aplica a solicitantes de trabajos de seguridad delicados. Las pruebas del polígrafo tienen fama de ser 90% precisas; esto es, atrapan a 90% de las personas que están mintiendo y validan a 90% de las personas que dicen la verdad. Pero, ¿cuántas personas que no pasan una prueba del polígrafo en realidad dicen la verdad? La mayoría de la gente cree que sólo 10% de esas personas que no pasan una prueba son identificadas falsamente. De hecho, el porcentaje real de acusaciones falsas puede ser mucho mayor. ¿Cómo es posible?

Pronto analizaremos la respuesta, pero la moraleja de esta historia ya debe ser clara. Aunque hagamos una descripción cuidadosa de datos, de acuerdo con los principios estudiados en las primeras tres secciones de este capítulo, aún podríamos ser llevados a conclusiones sorprendentes. Antes de pasar al tema del polígrafo, iniciamos con un par de sorpresas estadísticas.

Mejor en cada caso, pero peor de manera global

Suponga que una compañía farmacéutica crea un nuevo tratamiento para el acné. Para decidir su eficacia, la compañía proporciona el tratamiento actual a 90 pacientes y proporciona el nuevo tratamiento a 110 pacientes. Algunos pacientes tienen acné ligero y otros acné severo. La tabla 4.6 resume los resultados después de cuatro semanas de tratamiento, dividida en los tratamientos que fueron dados y si el acné fue ligero o severo. Si estudia con cuidado la tabla, notará estos hechos importantes:

- Entre los pacientes con acné *ligero*.
10 recibieron el tratamiento actual y 2 se curaron, para una tasa de curación de 20%.
90 recibieron el tratamiento nuevo y 30 se curaron, para una tasa de curación de 33%.
- Entre los pacientes con acné *severo*.
80 recibieron el tratamiento actual y 40 se curaron, para una tasa de curación de 50%.
20 recibieron el tratamiento nuevo y 12 se curaron, para una tasa de curación de 60%.

Tabla 4.6 Resultados del tratamiento para el acné

	Acné ligero		Acné severo	
	Curado	No curado	Curado	No curado
Tratamiento actual	2	8	40	40
Tratamiento nuevo	30	60	12	8

Observe que el nuevo tratamiento tiene una tasa más alta de curación *para ambos tipos* de pacientes, con acné ligero (33% para el nuevo tratamiento en comparación con 20% para el actual) y para pacientes con acné severo (60% para el nuevo tratamiento en comparación con 50% para el actual). Por tanto, ¿es justo para la compañía asegurar que su nuevo tratamiento es mejor que el tratamiento actual?

En principio, esto podría tener sentido. Pero, en lugar de revisar los datos para los pacientes con acné ligero y acné severo por separado, revisemos los resultados *globales*:

- Un total de 90 pacientes recibió el tratamiento actual y 42 se curaron (2 de 10 con acné ligero y 40 de 80 con acné severo), para una tasa global de curación de $42/90 = 46.7\%$.
- Un total de 110 pacientes recibieron el tratamiento nuevo y 42 se curaron (30 de 90 con acné ligero y 12 de 20 con acné severo), para una tasa global de curación de $42/110 = 38.2\%$.

Globalmente, el tratamiento *actual* tiene la tasa más alta de curación, a pesar de que el tratamiento nuevo tiene una tasa mayor para ambos casos de acné, ligero y severo.

Apropósito...

El caso general en que un conjunto de datos da resultados diferentes para cada uno de los diversos grupos de comparaciones que cuando los grupos son tomados en conjunto se conoce como *paradoja de Simpson*, llamada así porque fue descrita por Edward Simpson en 1951. Sin embargo, en realidad, la misma idea fue descrita alrededor de 1900 por el estadístico escocés George Yule.

Este ejemplo ilustra que es posible que algo parezca ser mejor en cada uno de dos o más grupos de comparaciones, pero que en términos globales sea peor. Si observa con cuidado, verá que esto ocurre a consecuencia de cómo se dividen los resultados globales entre grupos de diferentes tamaños (en este caso, los pacientes con acné ligero y los pacientes con acné severo).

EJEMPLO 1 ¿Quién jugó mejor?

La tabla 4.7 proporciona el desempeño en tiros de dos jugadores en cada medio de un juego de básquetbol. Shaq tuvo un porcentaje más alto de tiros en ambos, en el primer medio (40% a 25%) y en el segundo medio (75% a 70%). ¿Puede asegurar Shaq que él tuvo el mejor juego?

Tabla 4.7 Tiros de básquetbol

Jugador	Primer medio			Segundo medio		
	Canastas	Intentos	Porcentaje	Canastas	Intentos	Porcentaje
Shaq	4	10	40%	3	4	75%
Vince	1	4	25%	7	10	70%

Solución No, y podemos ver por qué revisando las estadísticas totales del juego. Shaq realizó un total de 7 canastas (4 en la primera mitad y 3 en la segunda) en 14 tiros (10 en la primera mitad y 4 en la segunda), para un porcentaje total de tiros de $7/14 = 50\%$. Vince realizó un total de 8 canastas en 14 tiros, para un porcentaje global de tiros de $8/14 = 57.1\%$. Sorpresa, aunque Shaq tiene un porcentaje de tiros más alto en ambos medios, Vince tiene un mejor porcentaje global para el juego.

¿Una mamografía positiva significa cáncer?

Con frecuencia asociamos tumores con cánceres, pero la mayoría de los tumores no son cáncer. En términos médicos, cualquier clase de hinchazón o crecimiento de tejido anormal es considerado un tumor. Un tumor causado por cáncer es *maligno* (o *canceroso*), todos los demás son *benignos*.

Imagine que usted es un doctor o una enfermera que trata a una paciente que tiene un tumor en el pecho. La paciente estará comprensiblemente nerviosa, pero usted puede darle un poco de consuelo diciéndole que sólo 1 de 100 tumores de pecho se vuelven malignos. Pero, para estar seguro, ordena una mamografía con el fin de determinar si su tumor es uno del 1% que son malignos.

Ahora, suponga que la mamografía reporta positivo, sugiriendo que el tumor es maligno. Las mamografías no son perfectas, por lo que el resultado positivo no necesariamente significa que su paciente tenga cáncer de pecho. Más en específico, suponga que el examen de la mamografía tiene 85% de precisión: identificará correctamente tumores malignos como malignos 85% de las veces y tumores benignos como benignos 85% de las veces. Cuando usted informa a su paciente que su mamografía fue positiva, ¿qué debe decirle respecto a su posibilidad de que tenga cáncer?

Puesto que el examen de la mamografía es 85% preciso, la mayoría de la gente supone que el resultados positivo significa que la paciente probablemente tenga cáncer. Los estudios han mostrado que la mayoría de los doctores creen que éste es el caso, y le dirían al paciente que se prepare para tratamiento contra el cáncer. Pero un análisis más cuidadoso muestra otra cosa. De hecho, las posibilidades de que el paciente tenga cáncer son muy pequeñas, alrededor de 5%. Podemos ver por qué analizando algunos números.

Considere un estudio en el que las mamografías se aplican a 10 000 mujeres con tumores en el pecho. Suponga que 1% de los tumores son malignos, $1\% \times 10\,000 = 100$ mujeres con cáncer; las restantes 9 900 mujeres tienen tumores benignos. La tabla 4.8 resume los resultados de la mamografía. Observe lo siguiente:

- El examen de la mamografía identifica correctamente 85% de los 100 tumores malignos como malignos. Así, da resultado positivo (maligno) para 85 de los tumores malignos; estos

A propósito...

Este ejemplo de mamografía y el ejemplo del polígrafo que sigue ilustra casos en que las probabilidades condicionales (analizadas en la sección 6.5) llevan a confusión. La manera apropiada de manejar las probabilidades condicionales fue descubierta por el reverendo Thomas Bayes (1702-1761) y con frecuencia se denomina *regla de Bayes*.

A propósito...

La precisión del examen de pecho rápidamente está mejorando; tecnologías más recientes, incluyendo mamografías digitales y ultrasonidos, parecen alcanzar precisiones de cerca de 98%. La prueba más definitiva para cáncer es una biopsia, aunque las biopsias podrían no acertar en la detección de cáncer si no se toman con suficiente cuidado. Si tiene exámenes negativos pero aún está preocupada por una anomalía consulte una segunda opinión. Podría salvar su vida.

Tabla 4.8 Resumen de resultados para 10 000 mamografías (cuando en realidad 100 tumores son malignos y 9 900 benignos)			
	El tumor es maligno	El tumor es benigno	Total
Mamografía positiva	85 verdaderos positivos	1 485 falsos positivos	1 570
Mamografía negativa	15 falsos negativos	8 415 verdaderos negativos	8 430
Total	100	9 900	10 000

casos se denominan **verdaderos positivos**. En los otros 15 casos malignos, el resultado es negativo, aunque en realidad las mujeres tengan cáncer, estos casos son **falsos negativos**.

- El examen de la mamografía identifica de manera correcta 85% de los 9900 tumores benignos como benignos. Así, da resultado negativo (benigno) para $85\% \times 9900 = 8415$ de los tumores benignos: estos casos son **verdaderos negativos**. Las restantes $9900 - 8415 = 1485$ mujeres dan resultados positivos en los que la mamografía de manera incorrecta identifica sus tumores como malignos; estos casos son **falsos positivos**.

En general, el examen de la mamografía da resultados positivos a 85 mujeres quienes en realidad tienen cáncer y a 1485 mujeres que *no* tienen cáncer. El número total de resultados positivos es $85 + 1485 = 1570$. Puesto que sólo 85 de éstas son verdaderos positivos (el resto son falsos positivos) la posibilidad de que un resultado positivo en realidad signifique cáncer es sólo $85/1570 = 0.054$, o 5.4%. Por tanto, cuando la mamografía de su paciente reporte positivo, debe asegurarle que sólo hay pocas posibilidades de que tenga cáncer.

EJEMPLO 2 Falsos negativos

Suponga que usted es un doctor que atiende a una paciente con un tumor en el pecho. Su mamografía reporta negativo. Con base en los números de la tabla 4.8, ¿cuál es la posibilidad de que ella tenga cáncer?

Solución Para los 10000 casos resumidos en la tabla 4.8, las mamografías son negativas para 15 mujeres que tienen cáncer y para 8415 mujeres con tumores benignos. El número total de resultados negativos es $15 + 8415 = 8430$. Así, la fracción de mujeres con cáncer que tienen falsos negativos es $15/8430 = 0.0018$, o ligeramente menos de 2 en 1000. En otras palabras, la posibilidad de que una mujer con una mamografía negativa tenga cáncer es de sólo 2 en 1000.

UN MOMENTO DE REFLEXIÓN

Aunque la posibilidad de cáncer con una mamografía negativa es pequeña, no es cero. Por tanto, podría ser buena idea hacer una biopsia a todos los tumores, sólo para asegurar. Sin embargo, las biopsias involucran cirugía, que son dolorosas y caras, entre otras cosas. Dados estos hechos, ¿usted cree que las biopsias deben ser rutinarias en todos los tumores? ¿Deben ser rutinarias para casos de mamografías positivas? Defienda su opinión.

Polígrafos y exámenes de drogas

Ahora estamos preparados para regresar a la pregunta realizada al inicio de esta sección, acerca de cómo una prueba de polígrafo 95% precisa puede llevar a un número sorprendente de acusaciones falsas. La explicación es muy similar a la que se utilizó en el caso de la mamografía.

Suponga que el gobierno proporciona una prueba de polígrafo a 1000 solicitantes para trabajos delicados de seguridad. Además suponga que 990 de estas 1000 personas dicen la

verdad en su prueba del polígrafo, mientras que sólo 10 mienten. Para una prueba que es 90% precisa, encontramos los resultados siguientes:

- De las 10 personas que mienten, el polígrafo identifica de manera correcta 90%, lo que significa que nueve no pasan la prueba (ellos se identifican como mentirosos) y una la aprueba.
- De las 990 personas que dicen la verdad, el polígrafo de manera correcta identifica 90%, lo que significa que $90\% \times 990 = 891$ personas veraces pasan la prueba y las otras $10\% \times 990 = 99$ personas veraces no pasan la prueba.

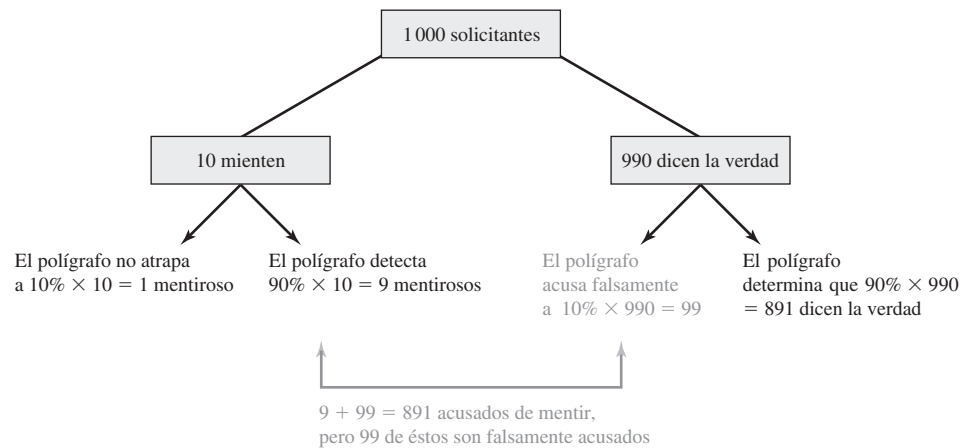


Figura 4.16 Un diagrama de árbol resume los resultados de una prueba de polígrafo 90% precisa para 1 000 personas, de las cuales sólo 10 mienten.

La figura 4.16 resume estos resultados. El número total de personas que no pasaron la prueba es $9 + 99 = 108$. De éstas, sólo 9 en realidad son mentirosas; las otras 99 fueron falsamente acusadas de mentir. Esto es, 99 de 108, o $99/108 = 91.7\%$, de las personas que no pasaron la prueba en realidad decían la verdad.

El porcentaje de personas que son acusadas falsamente en cualquier situación real depende de la precisión de la prueba y de la proporción de personas que están mintiendo. Sin embargo, para los números dados aquí, tenemos un resultado asombroso: suponiendo que el gobierno rechaza a los solicitantes que no pasan la prueba del polígrafo, entonces casi 92% de los solicitantes rechazados en realidad eran veraces, y podrían estar altamente calificados para los trabajos.

UN MOMENTO DE REFLEXIÓN

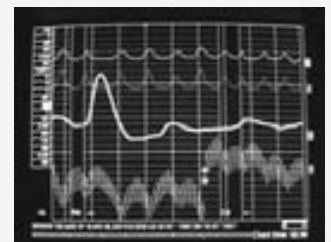
Imagine que usted es falsamente acusado de un crimen. La policía sugiere que si usted en realidad es inocente debe acceder a presentar una prueba en el polígrafo. ¿Lo haría? ¿Por qué sí o por qué no?

EJEMPLO 3 Prueba de droga en preuniversitarios

Todos los atletas que participaron en un campeonato preuniversitario regional de pista y campo deben proporcionar una muestra de orina para un examen de droga. Aquellos que no pasen la prueba son eliminados del encuentro y suspendidos de la competencia el año próximo. Los estudios muestran que en el laboratorio seleccionado las pruebas de droga son 95% precisos. Suponga que 4% de los atletas utilizan drogas. ¿Qué fracción de los atletas que no pasan la prueba son falsamente acusados y por tanto suspendidos sin culpa?

A propósito...

Un polígrafo, con frecuencia denominado detector de mentiras, mide una variedad de funciones corporales, incluyendo ritmo cardíaco, temperatura corporal y presión sanguínea. Los operadores del polígrafo buscan cambios sutiles en estas funciones que comúnmente ocurren cuando las personas mienten. Sin embargo, los resultados del polígrafo nunca han sido permitidos como evidencia en procesos criminales. Primero, 90% de precisión es demasiado bajo para la justicia. Además, estudios muestran que los polígrafos con facilidad son engañados por personas que se preparan para superarlos.



Solución La manera más sencilla de responder esta pregunta es utilizando algunos números muestrales. Suponga que hay 1000 atletas en el encuentro. Entonces 4%, o 40 atletas, en realidad utilizan drogas; el resto de los 960 atletas no emplean drogas. En ese caso, el examen de droga 95% preciso debe dar los resultados siguientes:

- 95% de los 40 atletas que utilizan drogas, o $0.95 \times 40 = 38$ atletas, no pasan el examen. Los otros dos atletas, que utilizan drogas, pasan el examen.
- 95% de los 960 atletas que no utilizan drogas pasan el examen, pero 5% de estos 960, o $0.05 \times 960 = 48$ atletas, no lo pasan.

El número total de atletas que no pasan el examen es $38 + 48 = 86$. Pero 48 de estos atletas que no pasaron el examen, o $48/86 = 56\%$, en realidad no usan drogas. A pesar del 95% de precisión del examen de droga más de la mitad de los estudiantes suspendidos son inocentes de uso de drogas.

Sección 4.4 Ejercicios

Alfabetización estadística y pensamiento crítico

1. **Falsos positivos y falsos negativos.** Cuando a alguien se le aplica un examen por uso de droga, ¿qué es un falso positivo? ¿Qué es un falso negativo?
2. **Resultado positivo en un examen.** Explique brevemente por qué un resultado positivo en una prueba de cáncer, tal como la mamografía, no necesariamente significa que el cliente tenga cáncer.
3. **Precisión de una prueba.** ¿Cómo es posible que una prueba del polígrafo pueda ser muy precisa mientras que tiene como resultado una gran proporción de acusaciones falsas?
4. **Mejor en cada medio, peor globalmente.** Cuando dos equipos de fútbol americano juegan, ¿un mariscal de campo puede tener un porcentaje más alto de pases en cada medio pero en todo el juego tener un porcentaje más bajo de pases?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Promedio de curso.** Ann y Bret están tomando el mismo curso de estadística, en el que la calificación final se determina mediante tareas y exámenes. El promedio de Ann en las tareas es mayor que el de Bret, y el promedio de Ann en el examen es mayor que el de Bret. Se concluye que el promedio global de Ann en el curso es mayor que el de Bret.
6. **Promedio de bateo.** El promedio de bateo de Ann en la primera mitad de la temporada es más alto que el de Bret, y el promedio de bateo de Ann en la segunda mitad de la temporada es mayor que el de Bret. Se concluye que el promedio de bateo de Ann en toda la temporada es mayor que el de Bret.

7. **Resultados de examen.** La probabilidad de tener inflamación si su prueba es positiva es la misma que la probabilidad de tener un examen positivo, si tiene inflamación.
8. **Precisión del examen.** Si un examen de droga es 90% preciso, 90% de aquéllos con examen positivo en realidad son usuarios de droga.

Conceptos y aplicaciones

9. **Porcentajes de bateo.** La tabla siguiente muestra los registros de bateo de dos jugadores de béisbol en la primera mitad (primeros 81 juegos) y la segunda mitad de una temporada.

Jugador	Primera mitad		
	Hits	Veces al bat	Promedio de bateo
Josh	50	150	.333
Jude	10	50	.200
Jugador	Segunda mitad		
	Hits	Veces al bat	Promedio de bateo
Josh	35	70	.500
Jude	70	150	.467

¿Quién tuvo el promedio de bateo más alto en la primera mitad de la temporada? ¿Quién tuvo el promedio más alto de bateo en la segunda mitad? ¿Quién tuvo el promedio de bateo global más alto? Explique cómo ilustran estos resultados la paradoja de Simpson.

10. **Porcentajes de pases.** La tabla siguiente muestra los registros de paso de dos mariscales de campo rivales en la primera mitad y en la segunda mitad de un juego de fútbol americano.

Jugador	Primera mitad		
	Completos	Intentos	Porcentaje
Allan	8	20	40%
Abner	2	6	33%
Jugador	Segunda mitad		
	Completos	Intentos	Porcentaje
Allan	3	6	50%
Abner	12	25	48%

¿Quién tuvo el promedio de completos más alto en la primera mitad? ¿Quién tuvo el promedio más alto de completos en la segunda mitad? ¿Quién tuvo el promedio de completos global más alto? Explique cómo ilustran estos resultados la paradoja de Simpson.

- 11. Calificaciones de examen.** La tabla siguiente muestra calificaciones de exámenes de matemáticas para alumnos de octavo grado en Nebraska y Nueva Jersey. Las calificaciones son separadas de acuerdo con la raza del estudiante. Además se muestran los promedios estatales para todas las razas.

	Blanco	No blanco	Promedio de todas las razas
Nebraska	281	250	277
Nueva Jersey	283	252	272

Fuente: Valoración Nacional de Progreso Educativo, de la revista *Chance*.

- ¿Cuál estado tenía las calificaciones más altas en ambas categorías raciales? ¿Cuál estado tuvo el promedio global más alto en ambas categorías raciales?
- Explique cómo un estado podría tener calificaciones más bajas en ambas categorías y aun así tener un promedio global más alto.
- Ahora considere la tabla siguiente, la cual proporciona los porcentajes de blancos y no blancos en cada estado. Utilice estos porcentajes para verificar que el promedio global de calificaciones en Nebraska es 277, como asegura la primera tabla.

	Blanco	No blanco
Nebraska	87%	13%
Nueva Jersey	66%	34%

- Utilice los porcentajes raciales para verificar que el promedio global de calificaciones en Nueva Jersey es 272, como se asegura en la primera tabla.
- Explique brevemente, con sus palabras, cómo apareció la paradoja de Simpson en este caso.

- 12. Calificaciones de exámenes.** Considere la tabla siguiente que compara la calificación promedio global (CPG) y las calificaciones del SAT de estudiantes de preparatoria en 1988 y 1998.

CPG	% estudiantes		Calificación del SAT		Cambio
	1988	1998	1988	1998	
A+	4	7	632	629	-3
A	11	15	586	582	-4
A-	13	16	556	554	-2
B	53	48	490	487	-3
C	19	14	431	428	-3
Promedio global			504	514	+10

Fuente: Citado en *Chance*, vol 12, número 2, 1999, de datos en el *New York Times*, 2 de septiembre de 1999.

- En términos generales, ¿cómo cambiaron las calificaciones del SAT de los estudiantes en las cinco categorías de calificación entre 1988 y 1998?
 - ¿Cómo cambió el promedio global de las calificaciones del SAT entre 1988 y 1998?
 - ¿Cómo es éste un ejemplo de la paradoja de Simpson?
- 13. Muertes por tuberculosis.** La tabla siguiente muestra las muertes debidas a la tuberculosis (TB) en la ciudad de Nueva York y en Richmond, Virginia, en 1910.

Raza	Nueva York	
	Población	Muertes por TB
Blanca	4 675 000	8 400
No blanca	92 000	500
Total	4 767 000	8 900
Raza	Richmond	
	Población	Muertes por TB
Blanca	81 000	130
No blanca	47 000	160
Total	128 000	290

Fuente: Cohen y Nagel, *An Introduction to Logic and Scientific Method*, Harcourt, Brace and World, 1934.

- Calcule la tasa de mortandad para blancos, no blancos y todos los residentes de la ciudad de Nueva York.
 - Calcule la tasa de mortandad para blancos, no blancos y residentes de Richmond.
 - Explique por qué éste es un ejemplo de la paradoja de Simpson y explique cómo surge la paradoja.
- 14. Gimnasia con pesas.** Dos equipos de corredores de campo travesía participaron en un estudio (hipotéticamente) en el que una parte de cada equipo usó gimnasia con pesas para complementar un ejercicio de carreras. El resto de los

corredores no utilizaron el entrenamiento con pesas. Al final de la temporada, la mejora media en los tiempos de carreras (en segundos) se registró en la tabla siguiente.

	Mejora media (segundos)		
	Entrenamiento con pesas	Entrenamiento sin pesas	Promedio del equipo
Gacelas	10	2	6.0
Guepardo	9	1	6.2

Describa cómo surge la paradoja en esta tabla. Resuelva la paradoja encontrando el porcentaje de cada equipo que utilizó el entrenamiento con pesas.

15. **Registros de básquetbol.** Considere los registros (hipotéticos) siguientes de básquetbol para los colegios Spelman y Morehouse.

	Colegio Spelman	Colegio Morehouse
Juegos en casa	10 ganados, 19 perdidos	9 ganados, 19 perdidos
Juegos fuera de casa	12 ganados, 4 perdidos	56 ganados, 20 perdidos

- a. Proporcione evidencia numérica para apoyar la afirmación de que el Colegio Spelman tiene un mejor equipo que el colegio Morehouse.
- b. Proporcione evidencia numérica para apoyar la afirmación de que el Colegio Morehouse tiene un mejor equipo que el colegio Spelman.
- c. ¿Cuál afirmación considera que tiene más sentido? ¿Por qué?

16. **Mejor medicamento.** Dos medicamentos, A y B, se probaron en un total de 2000 pacientes, de los cuales la mitad fueron mujeres y la mitad fueron hombres. El medicamento A se les dio a 900 pacientes y el medicamento B a 1100 pacientes. Los resultados aparecen en la tabla siguiente.

	Mujeres	Hombres
Medicamento A	5 de 100 curadas	400 de 800 curados
Medicamento B	101 de 900 curadas	196 de 200 curados

- a. Proporcione evidencia numérica para apoyar la afirmación de que el medicamento B es más efectivo que el medicamento A.
 - b. Proporcione evidencia numérica para apoyar la afirmación de que el medicamento A es más efectivo que el medicamento B.
 - c. ¿Cuál afirmación considera que tiene más sentido?
17. **Prueba del polígrafo.** Suponga que un polígrafo es 90% preciso (detectará de manera correcta 90% de las personas que están mintiendo y detectará correctamente 90% de las

personas que dicen la verdad). A 2000 empleados de una compañía se les aplica una prueba del polígrafo durante la cual se les pregunta si han usado drogas. Todos ellos niegan haberlas usado, de hecho 1% de los empleados en realidad usan drogas. Suponga que alguien a quien el operador del polígrafo detecta como embustero es acusado de mentiroso.

- a. Verifique que las entradas en la tabla siguiente se deducen de la información dada. Explique cada entrada.

	Usuarios	No usuarios	Total
La prueba determina que el empleado miente	18	198	216
La prueba determina que el empleado dice la verdad	2	1 782	1 784
Total	20	1 980	2 000

- b. ¿Cuántos empleados son acusados de mentirosos? De éstos, ¿cuántos en realidad mentían y cuántos decían la verdad?
- c. ¿Cuántos empleados se determinó que decían la verdad? De éstos, ¿cuántos en realidad decían la verdad? ¿Qué porcentaje de aquellos que se determinó que decían la verdad, en realidad eran veraces?

18. **Prueba de enfermedad.** Suponga que un examen para una enfermedad es 80% preciso para aquellos que tienen la enfermedad (verdaderos positivos) y 80% preciso para quienes no tienen la enfermedad (verdaderos negativos). En una muestra de 4 000 pacientes, la tasa de incidencia de la enfermedad coincide con el promedio nacional, que es 1.5%.

- a. Verifique que las entradas en la tabla siguiente se deducen de la información dada y que la incidencia global es 1.5%. Explique.

	Enfermedad	No enfermedad	Total
Examen positivo	48	788	836
Examen negativo	12	3 152	3 164
Total	60	3 940	4 000

- b. De aquéllos con la enfermedad, ¿qué porcentaje tuvo examen positivo?
- c. De aquéllos con examen positivo, ¿qué porcentaje tiene la enfermedad? Compare este resultado con el del inciso b y explique por qué son diferentes.
- d. Suponga que un paciente tiene un examen positivo para enfermedad. Como doctor que utiliza esta tabla, ¿cómo describiría las posibilidades del paciente de tener en realidad la enfermedad? Compare la cifra con la incidencia global de la enfermedad.

Más aplicaciones

19. **Estadísticas de contratación.** (Este problema tiene como base un ejemplo en “Ask Marilyn”, *Parade Magazine*,

28 de abril de 1996). Una compañía decidió expandirse, por lo que abrió una fábrica, lo que generó 455 empleos. Para los 70 puestos de oficina hubo 20 hombres y 200 mujeres solicitantes. De los hombres que solicitaron el empleo, 20% fueron contratados, mientras que sólo 15% de mujeres fueron contratadas. De los 400 hombres que solicitaron puesto de obrero, 75% fueron contratados, mientras que 85% de las 100 mujeres que solicitaron fueron contratadas. ¿Cómo al examinar los puestos administrativos y de obreros de manera separada sugieren una preferencia de contratación por mujeres? ¿Los datos globales apoyan la idea de que la compañía contrata preferentemente mujeres? Explique por qué éste es un ejemplo de paradoja de Simpson y cómo la paradoja puede resolverse.

20. Pruebas con medicamento. (Este problema tiene como base un ejemplo en “Ask Marilyn”, *Parade Magazine*, 28 de abril de 1996). Una compañía realiza dos pruebas de dos tratamientos para una enfermedad. El tratamiento A sana 20% de los casos (40 de 200) y el tratamiento B sana 15% de los casos (30 de 200). En la segunda prueba, el tratamiento A cura 85% de los casos (85 de 100) y el tratamiento B sana 75% de los casos (300 de 400). ¿Cuál tratamiento tiene la mejor tasa de curación, individualmente, en las dos pruebas? En forma global, ¿cuál tratamiento tiene la mejor tasa de curación? Explique por qué éste es un ejemplo de paradoja de Simpson y cómo la paradoja puede resolverse.

21. Riesgo de VIH. El departamento estatal de salud de Nueva York estima una tasa de 10% de VIH para la población en riesgo y una tasa de 0.3% para la población en general. Las pruebas para el VIH son 95% precisas en detectar tanto los verdaderos negativos como los verdaderos positivos. Una selección aleatoria y el examen de 5 000 personas en riesgo y 20 000 personas de la población general tiene como resultados la tabla siguiente.

	Población en riesgo	
	Prueba positiva	Prueba negativa
Infectado	475	25
No infectado	225	4 275
	Población general	
	Prueba positiva	Prueba negativa
Infectado	57	3
No infectado	997	18 943

- Verifique que las tasas de incidencia para las poblaciones, general y en riesgo, son 0.3% y 10%, respectivamente. También verifique que las tasas de detección para estas poblaciones son 95%.
- Considere la población en riesgo. De aquéllos con VIH, ¿qué porcentaje tienen prueba positiva? De aquellos que tienen prueba positiva, ¿qué porcentaje tienen VIH? Explique por qué estos dos porcentajes son diferentes.

- Suponga que un paciente en la categoría en riesgo tiene un examen positivo para la enfermedad. Como doctor que utiliza esta tabla, ¿cómo describiría las posibilidades del paciente de que en realidad tenga la enfermedad? Compare esta cifra con la tasa de incidencia global de la enfermedad.
- Considere a la población general. De aquéllos con VIH, ¿qué porcentaje tiene prueba positiva? De aquéllos con prueba positiva, ¿qué porcentaje tiene VIH? Explique por qué estos dos porcentajes son diferentes.
- Suponga que un paciente en la población general tiene un examen positivo para la enfermedad. Como doctor que utiliza esta tabla, ¿cómo describiría las posibilidades de que el paciente en realidad tenga la enfermedad? Compare esta cifra con la tasa de incidencia global de la enfermedad.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 4 en www.aw.com/bbt.

- Argumentos del polígrafo.** Visite sitios web dedicados a oponerse o apoyar el uso de pruebas del polígrafo. Resuma sus argumentos de ambas partes, en específico teniendo en cuenta el papel que desempeñan las tasas de falsos negativos en la discusión.
- Pruebas de dopaje.** Explore el tema de pruebas de dopaje en el trabajo o en competencias atléticas. Analice la legalidad de las pruebas de dopaje en estos contextos y la precisión de las pruebas que se realizan comúnmente.
- Detección de cáncer.** Investigue las recomendaciones concernientes a la detección de rutina para algunos tipos de cáncer (por ejemplo, cáncer de pecho, cáncer de próstata o cáncer de colon). Explique cómo se mide la precisión de las pruebas de detección.

EN LAS NOTICIAS

- Polígrafos.** Encuentre un artículo reciente en el que alguien o algún grupo proponga una prueba de polígrafo para determinar si una persona dice la verdad. En vista de lo que usted sabe acerca de las pruebas del polígrafo, ¿considera que los resultados serían significativos? ¿Por qué sí o por qué no?
- Exámenes de dopaje y atletas.** Encuentre un artículo que tenga que ver con pruebas de dopaje y atletas. Resuma cómo las pruebas son utilizadas y analice si son confiables.

Ejercicios de repaso del capítulo

1. Respecto a los pesos (en gramos) de chocolates M&M seleccionados aleatoriamente:

M&M rojos:

0.751	0.841	0.856	0.799	0.966	0.859	0.857
0.942	0.873	0.809	0.890	0.878	0.905	

M&M verdes:

0.925	0.914	0.881	0.865	0.865	1.015	0.876
0.809	0.865	0.848	0.940	0.833	0.845	0.852
0.778	0.814	0.791	0.810	0.881		

- Determine la media y la mediana para cada conjunto de datos.
 - Determine el rango y la desviación estándar para cada uno de los dos conjuntos de datos.
 - Proporcione el resumen de los cinco números y construya un diagrama de cajas para cada uno de los dos conjuntos de datos.
 - Aplique la regla empírica del rango para estimar la desviación estándar de cada uno de los dos conjuntos de datos. ¿Qué tan bien funciona la regla en cada caso? Analice brevemente por qué funciona o no funciona bien.
 - Con base en todos sus resultados, compare y analice los dos conjuntos de datos en términos de sus medidas de tendencia central y la variación. ¿Parece existir una diferencia entre los pesos de los M&M rojos y los M&M verdes?
2. Combine las dos muestras de pesos del ejercicio anterior y determine:
- El percentil para el peso de 0.845 gramos.
 - La moda.
3. a. ¿Cuál es la desviación estándar para una muestra de 50 valores, en la cual todos son iguales?
- b. ¿Cuál de los dos acumuladores para automóvil preferiría comprar y por qué?
- Uno tomado de una población con vida media de 48 meses y una desviación estándar de 2 meses.
 - Uno tomado de una población con vida media de 48 meses y una desviación estándar de 6 meses.
- c. Si un dato atípico se incluye en una muestra de 50 valores, ¿cuál es el efecto del dato atípico en la media?
- d. Si un dato atípico se incluye en una muestra de 50 valores, ¿cuál es el efecto del dato atípico en la mediana?
- e. Si un dato atípico se incluye en una muestra de 50 valores, ¿cuál es el efecto del dato atípico en el rango?
- f. Si un dato atípico se incluye en una muestra de 50 valores, ¿cuál es el efecto del dato atípico en la desviación estándar?

Cuestionario del capítulo

- Cuando suma los valores 3, 5, 8, 12 y 20 y luego divide entre el número de valores, el resultado es 9.6. ¿Qué término describe mejor este valor: promedio, media, mediana, moda o desviación estándar?
- Para medir la variación en una muestra, ¿cuál de las estadísticas siguientes debe utilizarse: promedio, media, mediana, moda o desviación estándar?
- Cuando se utiliza la regla empírica del rango para determinar el valor de la desviación estándar de una muestra de datos, ¿el resultado siempre es el valor real de la desviación estándar?
- Un conjunto de datos consiste en un conjunto de valores numéricos. ¿Cuál, si lo hay, de los enunciados siguientes podría ser correcto?
 - No existe moda
 - Existen dos modas.
 - Existen tres modas.
- Indique si el enunciado dado podría aplicarse a un conjunto de datos de 1 000 valores todos diferentes.
 - El percentil 20o. es mayor que el percentil 30o.
 - La mediana es mayor que el primer cuartil.
 - El tercer cuartil es mayor que el primer cuartil.
 - La media es igual a la mediana.
 - El rango es cero.
- Una prueba estándar para la ansiedad se diseña de manera que la calificación media es 50 y la desviación estándar es 10. Con base en la regla empírica del rango, ¿cuáles son los probables valores mínimo y máximo?
- Utilice la regla empírica del rango para estimar la desviación estándar de los valores 2, 4, 5, 8 y 10.
- ¿Cuál es la desviación estándar de los valores 5.8, 5.8, 5.8, 5.8 y 5.8?
- ¿Cuál es el rango de los valores 2.0, 3.7, 4.9, 5.0, 5.7, 6.7., 8.5 y 9.0?
- Identifique los componentes que constituyen el resumen de los cinco números para un conjunto de datos.

Uso de la tecnología

Muchos programas de cómputo le permiten introducir un conjunto de datos y utilizar una operación para obtener varias estadísticas muestrales diferentes, que hacen referencia a *estadísticas descriptivas*. Tales estadísticas incluyen, con frecuencia, a la media, mediana, rango y la desviación estándar. A continuación están algunos de los procedimientos para obtener tales estadísticas.

Observaciones importantes acerca de los procedimientos siguientes:

1. SPSS y Excel proporcionan diagramas de cajas “modificados” que muestran puntos especiales que representan datos atípicos, que son valores lejanos de casi todos los demás valores.
2. No existe un acuerdo universal sobre un procedimiento único para el cálculo de los cuartiles y diferentes programas arrojan resultados diferentes. Por ejemplo, si utiliza el conjunto de datos 1, 3, 6, 10, 15, 21, 28 y 36, obtendrá estos resultados:

	Q ₁	Q ₂	Q ₃
SPSS	3.75	12.5	26.25
Excel	5.25	12.5	22.75
STATDISK	4.5	12.5	24.5
Minitab	3.75	12.5	26.25
SAS	4.5	12.5	24.5
TI-83/84 Plus	4.5	12.5	24.5

SPSS

Estadísticas descriptivas: inicie introduciendo un conjunto de datos o abriendo un conjunto de datos existente. Dé clic en **Analyze**, seleccione **Descriptive Statistics** y luego seleccione **Frequency**. Dé doble clic en la variable o columna de datos a la izquierda. Ahora dé clic en el botón **Statistics** y proceda a dar clic en los cuadros correspondientes a la estadística que necesite. Puede seleccionar cuartiles, desviación estándar, rango, mínimo, máximo, media, mediana y otras. Dé clic en el botón **Continue**; luego dé clic en el botón **OK** para obtener los valores de las estadísticas seleccionadas.

Boxplot (diagrama de cajas): ingrese o abra un conjunto de datos en **Graphs** y seleccione **Boxplot**. En la ventana boxplot seleccione la opción **Summaries of separate variable** (resúmenes de variables separadas) y **Simple**, y luego dé clic en el botón **Define**. Ahora dé clic en la variable o columna a la izquierda, de clic en el botón a la mitad del cuadro, y luego dé clic en **OK**.

Excel

Estadística descriptiva: el complemento **Análisis de datos** debe estar instalado. Si el procedimiento siguiente no produce la ventana de análisis de datos, debe instalar el complemento **Análisis de datos**. Si está utilizando Excel 2003, dé clic en **Ayuda**, luego introduzca **Paquete análisis de datos** y presione **Enter**. Seleccione la opción de ayuda **Cargar y descargar complementos**, se mostrarán las instrucciones. Si está utilizando Excel 2007 seleccione la función de ayuda dando clic en el icono del símbolo de interrogación. Luego introduzca **Análisis de datos** y seleccione el tema de ayuda **Cargar el complemento de análisis** y aparecerán las instrucciones.

Introduzca los datos muestrales en la columna A. Seleccione **Herramientas** y luego **Análisis de datos**. Si utiliza Excel 2007 dé clic en **Datos** y luego en **Análisis de datos**. En la ventana que se muestra seleccione **Estadística descriptiva**, y dé clic en **Aceptar**. En el cuadro de diálogo introduzca el rango de entrada (tal como A1:A76 para 76 valores en la columna A), dé clic en **Resumen de estadísticas** y luego dé clic en **Aceptar**. Las estadísticas descriptivas incluyen la media, la mediana, la desviación estándar, el rango, el mínimo y el máximo.

Diagramas de cajas: Aunque Excel no está diseñado para generar diagramas de cajas, éstos pueden generarse usando el complemento DDXL que es un complemento de este libro.

Primero introduzca los datos en la columna A. Seleccione **Charts and Plots**. Para el tipo de función, seleccione la opción **Boxplot**. En el cuadro de diálogo dé clic en el icono del lápiz e introduzca el rango de datos, tal como A1:A76, si tiene 76 valores listados en la columna A. Dé clic en **OK**. El resultado será un diagrama modificado de cajas con datos atípicos atenuados y datos atípicos extremos. Además se mostrará el resumen de los cinco números.

STATDISK

Estadística descriptiva: ingrese los datos en la ventana de datos. Dé clic en **Data** y seleccione **Descriptive Statistics**. Ahora dé clic en **Evaluate** para obtener varias estadísticas descriptivas, incluyendo media, mediana, rango, desviación estándar y los valores que constituyen el resumen de los cinco números (mínimo, primer cuartil, mediana o segundo cuartil, tercer cuartil y máximo).

Diagrama de cajas: ingrese los datos en la ventana de datos, luego dé clic en **Data** y luego en **Boxplot**. Dé clic en las columnas que quiere incluir y luego dé clic en **Plot**.

HABLEMOS DE LA BOLSA DE VALORES



¿Qué es promedio cuando se habla del Dow Jones?

Cuando se trata de “promedios”, éste es extraordinario. No puede ver las noticias sin escuchar qué le está sucediendo, y cada día muchas personas dedican horas a su seguimiento. Es por mucho el más famoso indicador del desempeño de la bolsa de valores. Por supuesto estamos hablando del índice Dow Jones (*Dow Jones Industrial Average* o en breve DJIA). ¿Pero qué es exactamente este índice?

La forma más simple de entender este índice es investigar su historia. Cuando arrancó la moderna era industrial, a finales del siglo XIX, muchas personas consideraron a las acciones muy peligrosas e inversiones demasiado especulativas. Una de las razones fue la falta de control, lo que hizo fácil que especuladores ricos, gerentes sin escrúpulos y empresarios voraces manipularan los precios de las acciones. Pero otra razón fue que, dado lo complejo del comercio diario de acciones, incluso para los profesionales de Wall Street, tenían poco tiempo para calcular si las acciones en general estaban subiendo (un “mercado a la alza”) o estaban bajando (“mercado a la baja”). Charles H. Dow, fundador (junto con Edward D. Jones) y primer editor de *Wall Street Journal* creía que podía corregir este problema creando un “promedio” para la bolsa de valores

en su totalidad. Si el promedio era alto, el mercado tenía superávit y si el promedio era bajo el mercado tenía déficit.

Para mantener el promedio (ahora conocido como índice), Dow eligió 12 grandes empresas para incluirlas en su promedio. El 26 de mayo de 1896 Dow sumó los precios de las acciones de estas 12 empresas y las dividió entre 12, con lo que determinó un precio medio de las acciones de \$40.94. Éste fue el primer valor del DJIA. Como Dow había esperado, repentinamente fue fácil para el público seguir la dirección del mercado con sólo comparar su promedio día a día, mes con mes o año con año.

La idea básica que sustenta el DJIA sigue siendo la misma hoy en día, la única diferencia es que hoy se incluyen 30 acciones en lugar de 12 (vea la tabla 3.5). Sin embargo, el DJIA ya no es el precio medio de sus 30 acciones. En lugar de eso, se calcula sumando los precios de las 30 acciones y se divide entre un divisor especial. A consecuencia de este divisor, ahora se considera el DJIA como un índice que nos ayuda a dar un seguimiento de los valores de las acciones, en lugar de ser un promedio real de precios de acciones.

El divisor está diseñado para dar continuidad en el valor subyacente que representa el DJIA y, por tanto, debe cambiar siempre que la lista de 30 acciones cambie y siempre que una compañía en la lista tenga una división de acciones. Un ejemplo sencillo muestra por qué el divisor debe cambiar cuando la lista cambia. Suponga que el DJIA consiste en sólo dos acciones (en lugar de 30): la acción A con un precio de \$100 y la acción B con un precio de \$50. El precio medio de estas 2 acciones es $(\$100 + \$50)/2 = \$75$. Ahora supongamos que cambiamos de la lista la acción B y la reemplazamos con la acción C y que la acción C tiene un precio de \$200. La nueva media es $(\$100 + \$200)/2 = \$150$, por lo que si sólo reemplazamos una acción de la lista aumentaría el precio medio de \$75 a \$150. Por tanto, para mantener el “valor” del DJIA constante cuando cambia esta lista, debemos dividir la nueva media de \$150 entre 2. De esta forma, el DJIA permanece en 75 antes y después de que la lista cambie, pero ya no podemos pensar en este 75 como un precio medio en dólares.

Para ver por qué una división de acciones cambia el divisor, nuevamente supongamos que el índice consiste en sólo dos acciones: la acción X en \$100 y la acción Y con un valor de \$50, para un precio medio de \$75. Ahora, suponga que la acción X sufre una división de 2 por 1, por

tanto su nuevo valor será de \$50. Con ambas acciones ahora valuadas en \$50, el precio medio después de la división de la acción también sería \$50. En otras palabras, aunque una división de acciones no afecta el valor total de la compañía (sólo cambia el número y precios de los títulos) no deberíamos tener una caída en el precio de \$75 a \$50. En este caso, podemos preservar la continuidad dividiendo la nueva media de 50 entre $2/3$ (que es equivalente a multiplicar por $3/2$) por lo que el DJIA se mantiene en 75 antes y después de la división de la acción.

Al igual que en estos ejemplos sencillos, el divisor real cambia con cada cambio en la lista o al dividirse las acciones, por lo que éste ha cambiado muchas veces desde que Charles Dow lo calculó por primera vez como una media real. El valor actual del divisor se publica diariamente en el *Wall Street Journal*.

Dado que en la actualidad hay más de 10 000 acciones que se comercian de manera activa, podría parecer asombroso que una muestra de sólo 30 pueda reflejar la actividad global del mercado. Pero ahora, cuando las computadoras hacen más sencillo calcular “promedios” en el mercado de valores de muchas otras maneras, podemos revisar datos históricos y ver que el DJIA en verdad ha sido un indicador confiable del desempeño global del mercado. La figura 4.17 muestra el desempeño histórico del DJIA.

Si estudia la figura 4.17 cuidadosamente, podría intentar pensar que puede ver patrones, que le permitirían pronosticar valores precisos del mercado en el futuro. Desafortunadamente, nadie ha encontrado una manera de hacer pronósticos precisos y la mayoría de los economistas creen que tales pronósticos son imposibles.

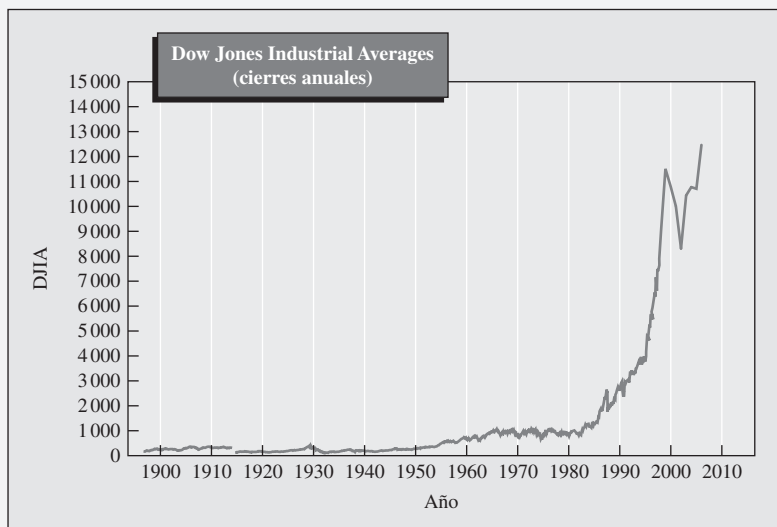


Figura 4.17 Valores al cierre del año del DJIA, hasta 2006.

La inutilidad de tratar de pronosticar el mercado se ilustra mediante la historia del estimado profesor Benjamin Graham, con frecuencia llamado el padre de “la investigación en valores”. En la primavera de 1951 uno de sus estudiantes fue por una asesoría de inversión. El profesor Graham notó que el DJIA se mantuvo en 250 pero caía por debajo de 200 al menos una vez cada año desde 1896. Puesto que en 1951 no había caído por debajo de los 200, el profesor Graham aconsejó a su estudiante que pospusiera la compra hasta que ésta bajara. El profesor Graham presumiblemente siguió su propio consejo, pero su estudiante no. En lugar de posponerlo de inmediato invirtió sus “10 000 dólares” en el mercado. Resultó que, en 1951, el mercado nunca cayó por debajo de los 200, ni en ningún otro momento desde entonces. Y el alumno, Warren Buffet, se convirtió en multimillonario.

PREGUNTAS PARA DISCUSIÓN

1. La bolsa de valores se sigue considerando un riesgo de inversión mayor que los bonos o las cuentas de ahorro. Sin embargo, los asesores financieros recomiendan, casi en forma

general, mantener al menos algunas acciones, lo que es un muy diferente a la situación que prevalecía hace más de un siglo. ¿Qué papel considera que desempeñó el DJIA en la conformación de la confianza de los inversionistas en la bolsa de valores?

2. El DJIA es uno de los tantos índices de la bolsa de valores hoy en día. Brevemente veremos otros índices como el S&P 500, el Russell 2000 y el NASDAQ. ¿Cómo es que estos índices difieren del DJIA? ¿Cree que alguno de estos índices debe considerarse más confiable del mercado global que el DJIA? ¿Por qué sí o por qué no?
3. Las 30 acciones en el DJIA representan una muestra de más de 10 000 acciones comercializadas activamente. Sin embargo no es una muestra aleatoria, porque son escogidas por editores específicos con motivos especiales dado que algunas veces se incluye el sesgo personal. Suponga que usted elige una muestra aleatoria de 30 acciones y hace un seguimiento de sus precios. ¿Considera que tal muestra aleatoria continuaría la trayectoria del mercado así como las acciones en el DJIA? ¿Por qué sí y por qué no?
4. Haga su propio “portafolio” con 10 acciones de su preferencia y suponga que posee 100 títulos de cada una. Calcule el valor total de su portafolio hoy y haga el seguimiento del precio hasta el siguiente mes. Al final del mes calcule el porcentaje de cambio de su portafolio. ¿Cómo compara el rendimiento de su portafolio con el rendimiento en el DJIA durante el mes? Si realmente fuera propietario de estas acciones ¿seguiría siéndolo o las vendería? Explique su respuesta.
5. La figura 4.17 muestra los valores al cierre del año del DJIA hasta 2006. Determine los datos de fin de año desde 2006 y agréguelos a la gráfica. Después determine los máximos y los mínimos de cada año desde 1995 y grafique cada uno de éstos como una curva separada en la gráfica. ¿Cómo compara los máximos y los mínimos al final de cada año? ¿Considera que los datos de cada año son una forma válida de hacer el seguimiento del cambio histórico en el DJIA?

LECTURAS RECOMENDADAS

Puede encontrar información del DJIA regularmente en el *Wall Street Journal*, *New York Times*, *Money*, *Business Week* y otras publicaciones periódicas.

Morris, K. M. y Siegel, A.M., *The Wall Street Journal Guide to Personal Finance*, Fireside, 2000.

Prestbo, John (ed.), *The Markets Measure: An Illustrated History of America Told Through the Dow Jones Industrial Average*, Dow Jones & Company, 1999.

HABLEMOS DE ECONOMÍA

¿Los ricos se vuelven más ricos?

En pleno auge de “la burbuja de las compañías punto com” en 1999, cuando los precios de las acciones de la tecnología se elevaron a niveles nunca antes vistos y que no se han visto desde entonces, el fundador de Microsoft, Bill Gates, se convirtió brevemente en la primera persona en el mundo con una fortuna de más de 100 mil millones de dólares. Su riqueza rebasó los productos nacionales brutos de todos los países, quitando a los 18 países más ricos del mundo. La riqueza de Gates, en realidad, ha disminuido desde entonces, debido a dos factores, el primero una disminución en el precio de las acciones de Microsoft y el segundo porque ha donado grandes cantidades de su riqueza neta a obras de caridad. No obstante, los medios de comunicación continuamente resaltan historias acerca de los súper ricos, haciendo parecer que los ricos se hacen cada vez más ricos mientras que el resto de nosotros nos estamos quedando atrás; pero, ¿esto es verdad?

Si queremos sacar conclusiones generales acerca de cómo el comportamiento económico de la persona promedio se compara con el de la persona rica, debemos revisar la distribución total del ingreso. Los economistas han desarrollado un número llamado índice Gini que es utilizado para describir el nivel de igualdad y de desigualdad de la distribución del ingreso. El índice Gini está definido de tal manera que sólo puede variar entre 0 y 1. Un índice Gini de 0 indica igualdad perfecta en el ingreso. Un índice Gini de 1 indica una perfecta desigualdad, en la cual una sola persona tiene todo el ingreso y todos los demás tienen nada. La figura 4.18 muestra el índice Gini en Estados Unidos desde 1947; observe que el índice Gini cayó de 1947 a 1968, esto indica que la distribución del ingreso se volvió más uniforme en este periodo. El índice Gini se ha elevado desde entonces e indica que los ricos, en efecto, se vuelven más ricos.

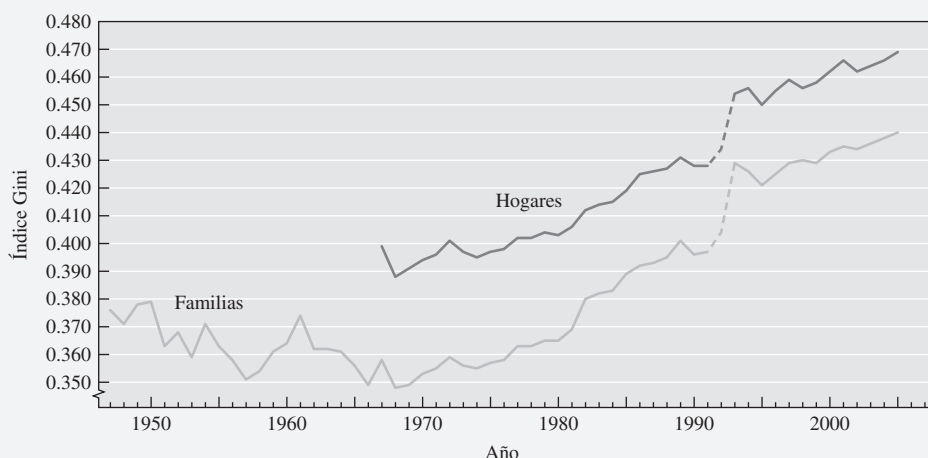


Figura 4.18 Índice Gini para familias y hogares, 1947-2005. La información de hogares, que incluye personas solas y hogares en los que los miembros no son parte de la misma familia, se han tomado desde 1967. Las líneas discontinuas en 1993 indican un cambio en la metodología para la recolección de información, por lo que el aumento correspondiente en el índice puede deberse en parte o completamente a este cambio, en lugar de que en realidad sea un cambio en la desigualdad del ingreso.

Fuente: Oficina del Censo, Estados Unidos.

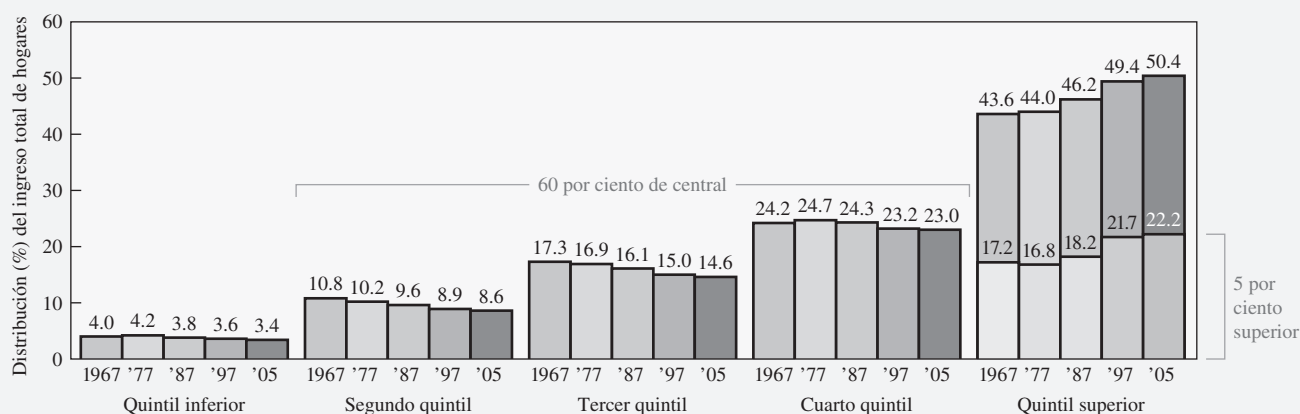


Figura 4.19 Distribución del ingreso total en hogares por quintiles (y el 5% superior): 1967, 1977, 1987, 1997 y 2005. Fuente: Oficina del Censo, Estados Unidos.

Aunque el índice Gini proporcione un sencillo resumen de un solo número de la desigualdad, este número en sí mismo es difícil de interpretar (y calcular). Una manera alterna de revisar la distribución del ingreso es estudiar los quintiles del ingreso, los cuales dividen a la población en cinco partes agrupadas por ingresos. Con frecuencia, el quintil más alto se divide para mostrar cómo el 5% superior de los que perciben ingresos se comparan con los demás.

La figura 4.19 muestra la parte del total de ingresos recibido por cada quintil y el 5% superior en diferentes décadas en Estados Unidos. La altura de cada columna (el número sobre ella) representa la parte del ingreso total. Por ejemplo, el 3.4 sobre el quintil más bajo en 2005 significa que el 20% de la población más pobre sólo recibió el 3.4% del ingreso total en Estados Unidos. De forma similar, el 50.4 sobre la columna del quintil superior significa que el 20% de la población más rica recibió 50.4% del ingreso. También observe que el 5% de los más ricos recibió el 22.2% del ingreso, casi el doble del total que recibió el 40% de la población más pobre. Si estudia esta gráfica cuidadosamente, usted verá que la parte del ingreso ganado por los primeros cuatro quintiles —lo que quiere decir todos excepto el 20% de la población más rica— cayó desde 1967. Mientras tanto, la parte ganada por el 20% de los más ricos se elevó de manera sustancial, así como lo hizo la parte del 5% superior. En otras palabras, esta gráfica también confirma que el rico se ha vuelto más rico comparado con la mayoría de la población.

Ahora que hemos establecido que el rico se vuelve más rico, la siguiente pregunta es si esto es importante. La mayoría de la gente, incluyendo muchos economistas, ha supuesto tradicionalmente que el aumento del ingreso desigual está mal debido a las democracias. Pero algunos economistas tanto de derecha como de izquierda del espectro político argumentan que el cambio en las recientes décadas es diferente. Por un lado, el cambio cumple una condición ética ampliamente aceptada, denominada criterio de Pareto, por el economista italiano Vilfredo Pareto (por quien los diagramas de Pareto se llaman así): cualquier cambio es bueno si ayuda sin perjudicar a alguien más. El criterio de Pareto parece satisfacerse ya que el crecimiento global en la economía de Estados Unidos ha ayudado a casi todos. En otras palabras, la mayoría de la gente tiene un porcentaje más pequeño de ingreso que lo que tuvo en el pasado, pero aún así tienen más ingreso absoluto y, por tanto, están viviendo mejor de lo que vivieron en el pasado.

En segundo lugar, el rico actual difiere del rico en el pasado. Por ejemplo, sin ir más lejos, en 1980 el 60% de los “Forbes 400” (las 400 personas más ricas) habían heredado la mayoría de sus fortunas. En 2005 menos del 20% de los 400 Forbes representaron dinero heredado. La implicación es que aunque usted hubiese nacido rico en el pasado, ahora podría *volverse* rico con buena educación y un trabajo duro. Seguramente ésta es una muy buena motivación para estudiar y trabajar duro.

Finalmente, aunque la desigualdad global en el ingreso ha aumentado, el ingreso desigual entre diferentes razas así como entre hombres y mujeres ha disminuido. En otras palabras, ahora es mucho más fácil que en el pasado para los afroamericanos, hispanos y mujeres poder ganar más que los hombres de raza blanca. Nuevamente, éste es un buen punto para los

valores democráticos estadounidenses, aunque todavía hay un largo camino por recorrer antes de que las desigualdades sean completamente eliminadas.

Por supuesto, aun si el reciente aumento en la desigualdad de ingresos ha sido bueno para Estados Unidos, no podría ser tan bueno si continúa. Desafortunadamente nadie sabe los motivos precisos que incrementan el ingreso desigual y por tanto nadie sabe cómo revertir la tendencia actual.

PREGUNTAS PARA DISCUSION

1. Compare diversas formas de ver la información mostrada en las figuras 4.18 y 4.19. Por ejemplo, ¿algo parece indicar un cambio más grande en el ingreso desigual que otro? ¿Puede usted considerar otras posibles formas de exhibir la información de ingresos que puedan dar una visión diferente de la que se mostró aquí?
2. ¿Está usted de acuerdo en que el criterio de Pareto es una buena manera de evaluar la ética de los cambios en la economía? ¿Por qué sí o por qué no?
3. Globalmente, ¿cree usted que el aumento en la desigualdad del ingreso ha sido bueno o malo para Estados Unidos?, ¿seguirá siendo bueno si esta tendencia continúa? Defienda su opinión.
4. Aunque la información económica sugiere que la gran mayoría de los estadounidenses está mejor hoy que como lo estuvieron en décadas pasadas, los estadounidenses más pobres aún viven en malas condiciones económicas. ¿Qué cree usted que se pueda o debería hacer para ayudar a mejorar la vida de los pobres? ¿Sus sugerencias pueden ser implementadas sin dañar la economía en general? Explique.

LECTURAS RECOMENDADAS

Usted puede encontrar extensa información y reportes sobre la distribución del ingreso si va al área de “Income” (ingresos) en el sitio web, www.census.gov, de la Oficina de Censos de Estados Unidos.

Arrow, Kenneth, *et al.* (eds.), *Meritocracy and Economic Equality*, Princeton University Press, 2000.

Firebaugh, Glenn, *The New Geography of Global Income Inequality*, Harvard University Press, 2006.

Nasar, Sylvia, “Is the U.S. Income Gap Really a Big Problem?”, *New York Times*, 4 de abril de 1999.



Nada en la vida debe temerse. Sólo debe entenderse.

—Marie Curie

Un mundo normal

CUANDO CAMINA EN UNA TIENDA, ¿CÓMO SABE SI UN precio de oferta en realidad es un buen precio? Cuando usted hace ejercicio y su ritmo cardíaco sube, ¿cómo sabe si ha subido suficiente, pero no demasiado, para ser resultado de un buen ejercicio? Si su hija de doce años corre una milla en 5 minutos, ¿ella es una futura promesa olímpica? Estas preguntas parecen muy diferentes, pero desde un punto de vista estadístico son muy similares. Cada una pregunta si un número particular (precio, ritmo cardíaco, tiempo de carrera) es, de alguna manera, poco usual. En este capítulo analizaremos cómo podemos responder tales preguntas con ayuda de la *distribución normal* que tiene forma de campana.

OBJETIVOS DE APRENDIZAJE

5.1 ¿Qué es normal?

Entender lo que significa una distribución normal y ser capaz de identificar situaciones en las que una distribución normal es muy probable que surja.

5.2 Propiedades de la distribución normal

Saber cómo interpretar la distribución normal en términos de la regla 68-95-99.7, calificaciones estándar y percentiles.

5.3 El teorema del límite central

Entender la idea básica subyacente al teorema del límite central y su importante papel en estadística.

5.1 ¿Qué es normal?

Suponga que una amiga está encinta y dará a luz el 30 de junio. ¿Usted le aconsejaría programar una importante reunión de negocios para el 16 de junio, dos semanas antes de su parto? La respuesta a esta pregunta requiere conocer si es probable que el bebé llegue con una anticipación de 14 días antes de la fecha del parto. Para eso necesitamos examinar la información concerniente a las fechas programadas y a las fechas reales de nacimiento.

La figura 5.1 es un histograma para una distribución de 300 nacimientos naturales en el hospital Providence Memorial; las fechas son hipotéticas, pero con base en cómo estarían distribuidos los nacimientos sin intervención médica. El eje horizontal muestra cuántos días antes o después de la fecha programada nació un bebé: números negativos representan nacimientos *antes* de la fecha programada, cero representa un nacimiento en la fecha programada y números positivos representan nacimientos *después* de la fecha programada. El eje vertical a la izquierda muestra el número de nacimientos para cada clase de 4 días. Por ejemplo, la frecuencia de 35 para la barra más alta corresponde a la clase de -2 a 2 días; muestra que del total de 300 nacimientos en la muestra, 35 nacimientos ocurrieron dentro de dos días de la fecha programada.

Para responder nuestra pregunta acerca de si es probable que un nacimiento ocurra con más de 14 días de anticipación, es más útil revisar las *frecuencias relativas*. Recuerde que la frecuencia relativa de cualquier valor es su frecuencia dividida entre el número total de valores (vea la sección 3.1). La figura 5.1 muestra las frecuencias relativas en el eje vertical de la derecha. Por ejemplo, la clase para -14 a -10 días (sombreado con gris oscuro), tiene una frecuencia relativa de alrededor de 0.07, o 7%. Esto es, alrededor de 7% de 300 nacimientos ocurrieron de 14 a 10 días antes de la fecha programada.

Ahora podemos determinar la proporción de nacimientos que ocurrieron más de 14 días antes de la fecha programada. Simplemente sumamos las frecuencias relativas para las clases a la izquierda de -14 , puede medir la gráfica para confirmar que estas clases tienen una frecuencia relativa total de alrededor de 0.21, lo cual indica que alrededor de 21% de los nacimientos en este conjunto de datos ocurrió más de 14 días antes de la fecha programada. Podríamos decir que su amiga tiene alrededor de 1 en 5 oportunidades de que su bebé nazca en o antes de la fecha de la junta de negocios. Si la reunión es importante, podría ser bueno programarla antes.

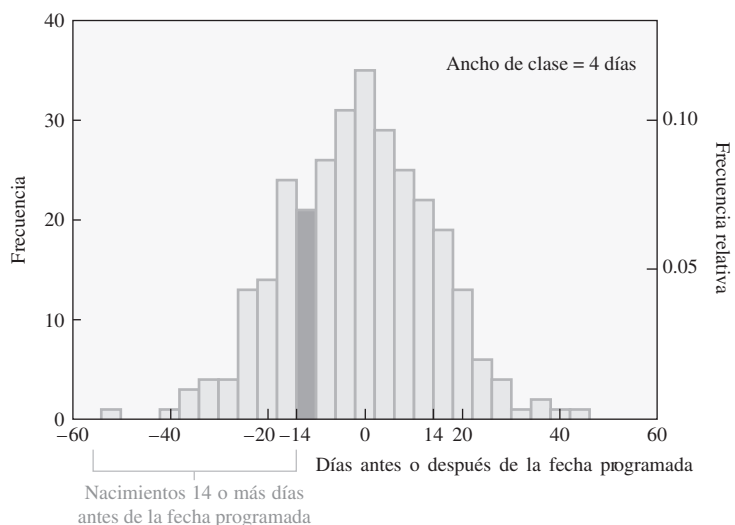


Figura 5.1 Histograma de frecuencias (eje izquierdo) y frecuencias relativas (eje derecho) para fechas de nacimiento relativas a la fecha programada. (Estos datos son hipotéticos). Los números negativos se refieren a nacimientos antes de la fecha programada, los números positivos se refieren a nacimientos posteriores a la fecha programada. El ancho de cada clase es 4 días. Por ejemplo, la clase sombreada en gris oscuro representa nacimientos que ocurrieron entre 10 y 14 días prematuramente.

UN MOMENTO DE REFLEXIÓN

Suponga que su amiga planea tomar tres meses de permiso por maternidad después de la fecha del nacimiento. Con base en los datos de la figura 5.1 y suponiendo una fecha programada para el 30 de junio, ¿ella se podría comprometer a estar en el trabajo el 10 de octubre?

La forma normal

La distribución de la fecha de nacimiento tiene una forma muy distintiva, la cual es más fácil de observar si sobreponemos en el histograma una curva continua (figura 5.2). Para nuestros propósitos presentes, la forma de esta distribución continua tiene tres características importantes:

- La distribución tiene *un solo pico*. Su moda, o fecha de nacimiento más común, es la fecha programada.
- La distribución es *simétrica* alrededor de su pico único: por tanto, su mediana y su media son iguales a la moda. La mediana es la fecha programada, ya que igual número de nacimientos ocurren antes y después de esta fecha. La media también es la fecha programada, ya que, para todo nacimiento previo a la fecha programada, existe un nacimiento posterior al mismo número de días a la fecha programada.
- La distribución tiene una dispersión tal que parece la forma de una campana, de modo que la denominamos distribución “en forma de campana”.

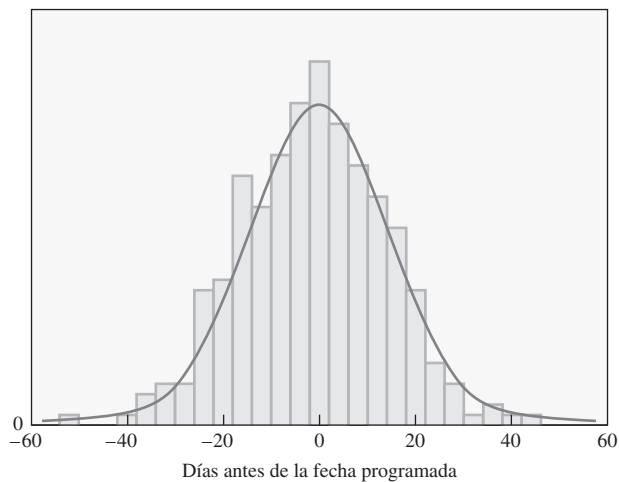


Figura 5.2 Una curva continua de la distribución normal se dibuja sobre el histograma de la figura 5.1.

UN MOMENTO DE REFLEXIÓN

El histograma de la figura 5.2, que tiene como base los nacimientos naturales, es casi simétrico. Ahora, los doctores por lo común inducen el nacimiento si una mujer se pasa demasiado de su fecha programada. ¿Cómo cambiaría la forma del histograma si se incluyen los nacimientos inducidos?

La distribución continua en la figura 5.2 con estas tres características se denomina **distribución normal**. (Observe que la fecha de nacimiento dada no es exactamente normal). Todas las distribuciones normales tienen la misma forma de campana característica, pero pueden diferir en su media y su variación. La figura 5.3 muestra dos distribuciones normales diferentes. Ambas tienen la misma media, pero la distribución (a) tiene mayor variación. Como analizaremos en la sección siguiente, el conocimiento de la *desviación estándar* (vea la sección 4.3) de una distribución normal nos dice todo lo que necesitamos saber acerca de su variación. Por tanto, una distribución normal puede describirse completamente con sólo dos números: su media y su desviación estándar.

NOTA TÉCNICA

Una distribución normal con media μ y desviación estándar σ se da mediante la fórmula

$$y = \frac{e^{-\frac{1}{2}[(x-\mu)/\sigma]^2}}{\sigma\sqrt{2\pi}}$$

Esta fórmula no se utiliza en este libro, pero describe de manera algebraica la forma de la distribución normal.

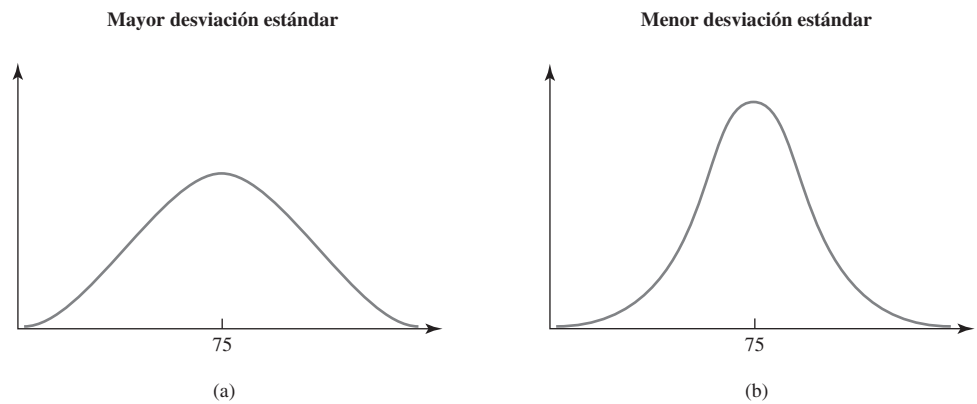


Figura 5.3 Ambas distribuciones son normales y tienen la misma media de 75, pero la distribución de la izquierda tiene una desviación estándar mayor.

Definición

La **distribución normal** es una distribución simétrica en forma de campana con un solo pico. Su pico corresponde a la media, mediana y moda de la distribución. Su variación puede caracterizarse mediante la desviación estándar de la distribución.

EJEMPLO 1 ¿Distribuciones normales?

La figura 5.4 muestra dos distribuciones: (a) un famoso conjunto de datos de las medidas de pechos de 5738 militares escoceses recolectados alrededor de 1846, y (b) la distribución de las densidades de población de los 50 estados de Estados Unidos. ¿Alguna de ellas es una distribución normal? Explique.

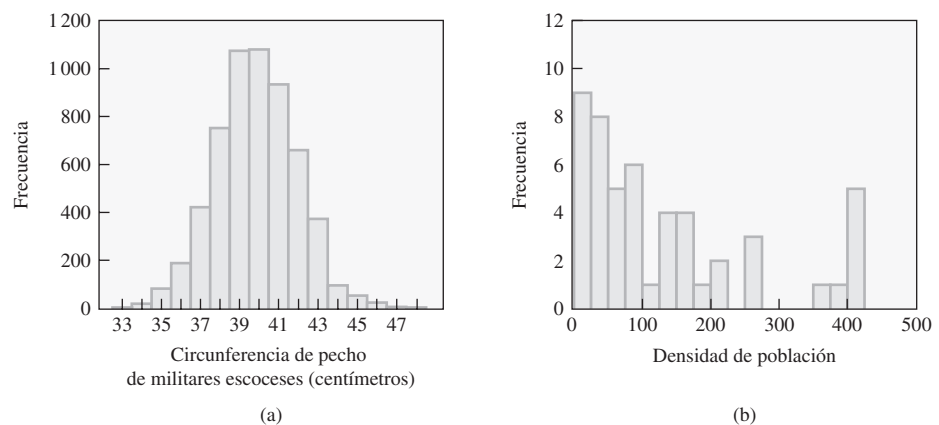


Figura 5.4 Fuente de (a) Adolphe Quetelet, *Lettres à S. A. R. le Duc Régnant de Saxe-Cobourg et Gotha*, 1846.

Solución La distribución en la figura 5.4a es casi simétrica, con una media entre 39 y 40 pulgadas; es casi normal. La distribución en la figura 5.4b muestra que la mayoría de los estados tienen densidades poblacionales bajas, y unas cuantas tienen densidades mucho más altas. Este hecho hace que la distribución sea sesgada a la derecha, por lo que no es una distribución normal.

La distribución normal y las frecuencias relativas

Recuerde que la frecuencia relativa total para cualquier conjunto de datos debe ser 1 (vea la sección 3.1). Ahora considere la curva continua para la distribución normal en la figura 5.2. Aunque ya no tenemos barras individuales, aún podemos asociar la altura de la curva normal

Apropósito...

Con los datos tomados de soldados franceses y escoceses, el científico social belga Adolphe Quetelet se dio cuenta en 1830 que las características humanas tales como estatura y circunferencia del pecho se distribuyen de manera normal. Esta observación lo llevó a acuñar el término “el hombre promedio”. Quetelet fue primer miembro extranjero de la Asociación Americana de Estadística (Estados Unidos).



con la frecuencia relativa. El hecho de que las frecuencias relativas deben sumar 1 condiciona que el área debajo de la curva normal debe ser 1.

La figura 5.5 proporciona un ejemplo de cómo interpretamos la curva normal. La idea clave es: *la frecuencia relativa para cualquier rango de datos es el área debajo de la curva que cubre ese rango de valores*. Por ejemplo, un cálculo preciso muestra que la región sombreada a la izquierda de -14 días comprende alrededor de 18% del área total bajo la curva. Por tanto, concluimos que la frecuencia relativa es alrededor de 0.18 para los valores menores de -14 días, lo que significa que alrededor de 18% de los nacimientos son prematuros en más de 14 días. De manera similar, la región sombreada a la derecha de 18 días incluye alrededor de 12% del área total bajo la curva. Por tanto, concluimos que la frecuencia relativa es alrededor de 0.12 para valores mayores de 18 días, lo que significa que alrededor de 12% de los nacimientos se retrasan más de 18 días. Juntos, vemos que $18\% + 12\% = 30\%$ de todos los nacimientos o son prematuros por más de 14 días o se retrasan por más de 18 días. (Observe que estas frecuencias relativas son correctas para la distribución normal continua de la figura 5.5, pero sólo son aproximaciones para los datos originales en el histograma de la figura 5.1).

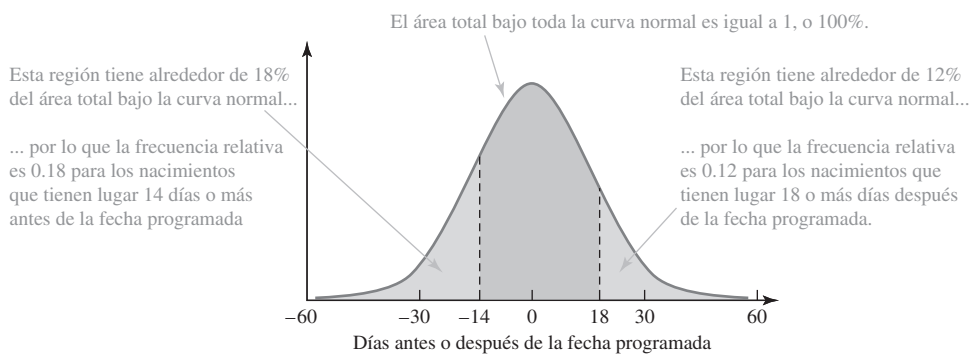


Figura 5.5 El porcentaje del área total en cualquier región bajo la curva normal nos dice la frecuencia relativa de los datos en ese grupo.

Frecuencias relativas y la distribución normal

- El área total que está debajo de la curva normal que corresponde a un rango de valores en el eje horizontal es la frecuencia relativa de esos valores.
- Debido a que la frecuencia relativa total debe ser de 1, el área total bajo la curva de distribución normal debe ser igual a 1, o el 100%.

UN MOMENTO DE REFLEXIÓN

De acuerdo con la figura 5.5, ¿qué porcentaje de nacimientos ocurren entre 14 días antes y 18 días después? Explique. (Sugerencia: recuerde que el área total debajo de la curva es 100%).

EJEMPLO 2 Estimación de áreas

Otra vez revise la distribución normal de la figura 5.5.

- Estime el porcentaje de nacimientos que ocurren entre 0 y 60 días después de la fecha programada.
- Estime el porcentaje de nacimientos que ocurren entre 14 días antes y 14 días después de la fecha programada.

Solución

- Alrededor de la mitad del área total debajo de la curva está en la región entre 0 y 60 días. Esto significa que alrededor de 50% de los nacimientos en la muestra ocurren entre 0 y 60 días después de la fecha programada.

A propósito...

El político escocés John Sinclair (1754-1835) fue uno de los primeros en recopilar datos de economía, demografía y agricultura. A él se le acredita la introducción de las palabras *statistics* y *statistical* en el idioma inglés, habiéndolos escuchado usar en alemán para referirse a cosas de estado (*state*, en inglés).

- b. La figura 5.5 muestra que alrededor de 18% de los nacimientos ocurren con más de 14 días de retraso respecto a la fecha programada. Puesto que la distribución es simétrica, alrededor de 18% de los nacimientos deben retrasarse por más de 14 días respecto a la fecha programada. Por tanto, un total de alrededor de $18\% + 18\% = 36\%$ de nacimientos ocurren 14 días antes o después de la fecha programada. La pregunta acerca de la región restante significa *entre* 14 días antes y 14 días después de la fecha programada, de modo que esta región debe representar $100\% - 36\% = 64\%$ de los nacimientos.

¿Cuándo podemos esperar una distribución normal?

La distribución normal es una buena aproximación a la distribución de muchas variables de interés práctico. Las características físicas tales como peso, estatura, presión arterial y tiempos de reflejo siguen una distribución normal. Las calificaciones de exámenes estandarizados tales como los exámenes SAT o del CI, se distribuyen de manera normal. Las estadísticas de deportes tales como promedios de bateo, tiempos en un evento de natación o resultados de salto de altura en una competencia de campo, también tienden a seguir una distribución normal. En realidad, gran parte de la estadística tiene como base la distribución normal. Sin embargo, no todas las variables siguen una distribución normal, por lo que es importante entender cuándo la distribución normal puede usarse y cuándo no es apropiada.

Al estudiar gráficas de distribuciones normales, podemos ver que conforman una distribución normal cuando:

- Tienen valores agrupados cerca de la media, de modo que tengan un solo pico, o sea unimodal.
- Los valores deben estar esparcidos de manera equitativa alrededor de la media, de modo que sean simétricos.
- Desviaciones grandes de la media deben ser cada vez más raras, por lo que tienen la característica forma de campana.

A un nivel de mayor profundidad, cualquier cantidad que esté influida por muchos factores diferentes tiende a estar distribuida de manera normal. Los rasgos físicos están influidos por muchos factores diferentes tanto genéticos como del ambiente. Las calificaciones de exámenes estandarizados reflejan el desempeño de muchos individuos trabajando en muchas preguntas diferentes del examen. Las estadísticas deportivas involucran muchos jugadores con diferentes desempeños en sus habilidades bajo diferentes condiciones.

A propósito...

La curva de la distribución normal con frecuencia se denomina *curva gaussiana* en honor del matemático alemán del siglo XIX Carl Friedrich Gauss. El lógico estadounidense Charles Pierce introdujo el término *distribución normal* alrededor de 1870.

Condiciones para una distribución normal

Un conjunto de datos que satisface los cuatro criterios siguientes es probable que tenga una distribución cercana a la normal.

1. La mayoría de los valores están agrupados cerca de la media, dando a la distribución un solo pico bien definido.
2. Los datos están dispersos de manera equitativa alrededor de la media, haciendo que la distribución sea simétrica.
3. Las desviaciones grandes de la media se hacen cada vez más raras, lo que produce delgadas colas de la distribución.
4. Los valores individuales resultan de una combinación de muchos factores diferentes, como factores genéticos y del ambiente.

EJEMPLO 3 ¿Es una distribución normal?

¿Cuál de las variables siguientes esperaría que tenga una distribución normal o cercana a la distribución normal?

- a. Calificaciones en un examen mucho muy sencillo.
- b. Estaturas de una muestra aleatoria de mujeres adultas.
- c. El número de manzanas Macintosh en cada una de 100 canastas llenas de un bushel (que equivale a 36.23 litros).

Solución

- a. Los exámenes tienen una calificación máxima (100%) que limita el tamaño de los valores. Si el examen es sencillo, la media será alta y muchas calificaciones serán cercanas a la máxima posible. Las pocas calificaciones bajas podrían distribuirse un poco abajo de la media. Por tanto, esperaríamos que la distribución de calificaciones sea sesgada a la izquierda y no normal.
- b. La estatura está determinada por una combinación de muchos factores (la conformación genética de ambos padres y quizá factores ambientales y nutricionales). Esperamos que la altura media para la muestra sea cercana a la moda (la altura más común). También esperamos que haya aproximadamente el mismo número de mujeres arriba que abajo de la media, y estaturas extremadamente grandes o pequeñas deben ser raras. Esto es por lo que la estatura se distribuye casi normalmente.
- c. El número de manzanas en una canasta de un bushel varía con el tamaño de las manzanas. Esperamos que en la distribución haya una sola moda que debe ser cercana al número medio de manzanas por canasta. El número de canastas con más que el número medio de manzanas debe ser cercano al número de canastas con menos manzanas que el número medio de manzanas. Por tanto, esperamos que el número de manzanas por canasta tenga una distribución casi normal.

UN MOMENTO DE REFLEXIÓN

¿Esperaríamos que las calificaciones se distribuyesen de manera normal en un examen moderadamente difícil? Sugiera dos o más cantidades que esperaría que se distribuyesen normalmente.

Sección 5.1 Ejercicios**Alfabetización estadística y pensamiento crítico**

1. **Distribución normal.** ¿Cuando nos referimos a distribución “normal”, la palabra *normal* tiene el mismo significado que en el lenguaje común, o tiene un significado especial en estadística? Exactamente, ¿qué es una distribución normal?
2. **Distribución normal.** Una distribución normal, de manera informal y vaga, se describe como una distribución de probabilidad que tiene “forma de campana” cuando se grafica. ¿Qué es la forma de “campana”?
3. **Dígitos aleatorios.** Con frecuencia las computadoras se utilizan para generar dígitos aleatorios de números telefónicos para llamar cuando se realiza una encuesta. ¿La distribución de esos dígitos es una distribución normal? ¿Por qué sí o por qué no?
4. **Áreas.** Si usted determina que, en la gráfica de una distribución normal, el área a la izquierda de un valor particular es 0.2, ¿cuál es el área a la derecha de ese valor?

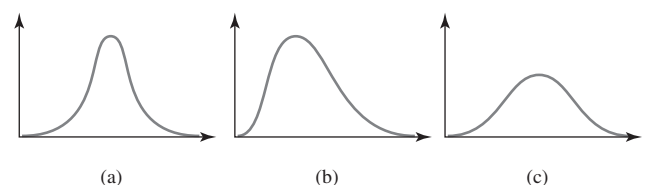
¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Calificaciones del SAT.** Una muestra aleatoria de calificaciones del SAT tiene una distribución normal con media de 1518 y una mediana de 1490.

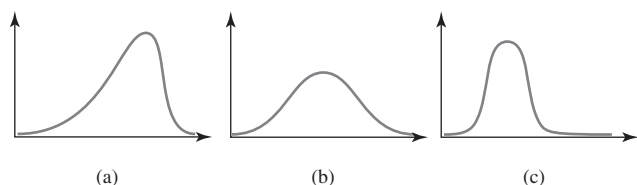
6. **Pesos.** Como parte de un estudio nacional de salud, cada persona en una muestra aleatoria de sujetos se pesa, y esos pesos tienen una distribución normal con dos modas.
7. **Calificaciones del CI.** La media de un conjunto distribuido normalmente de calificaciones de CI es 100, y 60% de las calificaciones son mayores a 105.
8. **Volúmenes de Pepsi.** Un analista de control de calidad mide los contenidos de 100 latas de Pepsi regular y pronostica que la distribución de volúmenes será una distribución normal.

Conceptos y aplicaciones

9. **¿Qué es normal?** Identifique la distribución en la figura 5.6 que no es normal. De las dos distribuciones normales, ¿cuál tiene mayor desviación estándar?

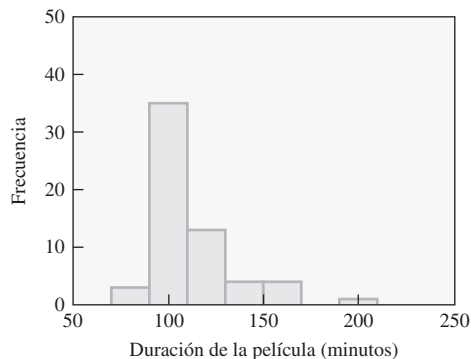
**Figura 5.6**

10. **¿Qué es normal?** Identifique la distribución en la figura 5.7 (en la página siguiente) que no es normal. De las dos distribuciones normales, ¿cuál tiene la mayor desviación estándar?

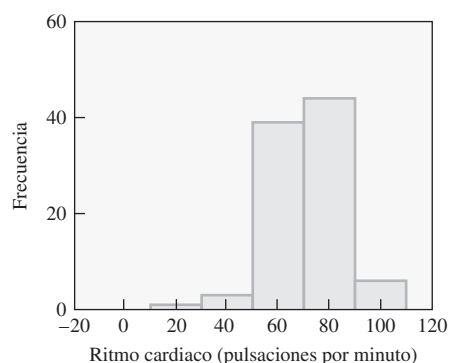
**Figura 5.7**

Variables normales. Para cada conjunto de datos en los ejercicios 11 al 18 indique si esperaría que se distribuya normalmente. Explique su razonamiento.

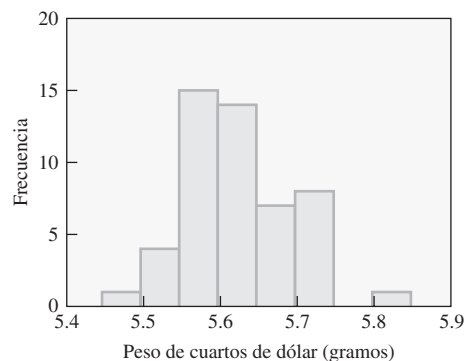
11. **Pesos de CD.** Los pesos exactos de una muestra aleatoria de CD Sony.
12. **Ingresos.** Los ingresos de adultos seleccionados de manera aleatoria en Estados Unidos.
13. **Tiros de un dado.** Los resultados de 500 tiros de un dado no cargado.
14. **Calificaciones del SAT.** Todas las calificaciones del SAT del año pasado.
15. **Fuerza de sujeción.** Las medidas de la fuerza de sujeción para una muestra aleatoria de hombres.
16. **Retraso en vuelos.** La duración de tiempo que las aeronaves comerciales se retrasan antes de despegar.
17. **Tiempos de espera.** Los tiempos de espera en una parada de autobús, si éste llega una vez cada diez minutos y usted llega en tiempos aleatorios.
18. **Multas.** Los montos de las multas por estacionarse en lugar prohibido en una muestra de 1000 automóviles.
19. **Duración de películas.** La figura 5.8 muestra un histograma para la duración de 60 películas. La duración de una película media es 110.5 minutos. ¿Esta distribución es cercana a la normal? ¿Esta variable debe tener una distribución normal? ¿Por qué sí o por qué no?

**Figura 5.8**

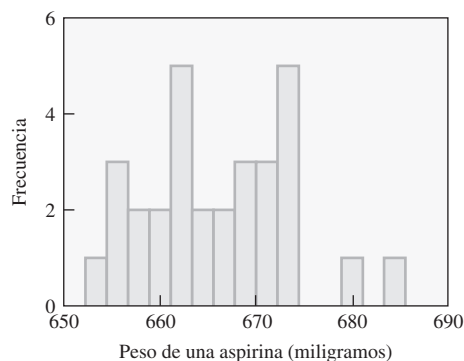
20. **Ritmos cardiacos.** La figura 5.9 muestra un histograma para los ritmos cardiacos de 98 estudiantes. El ritmo cardiaco medio es 71.2 pulsaciones por minuto. ¿Esta distribución es cercana a la normal? ¿Esta variable debe tener una distribución normal? ¿Por qué sí o por qué no?

**Figura 5.9**

21. **Pesos de monedas de 25 centavos.** La figura 5.10 muestra un histograma para los pesos de 50 monedas de 25 centavos elegidas de manera aleatoria. El peso medio es 5.62 gramos. ¿Esta distribución es cercana a la normal? ¿Esta variable debe tener una distribución normal? ¿Por qué sí o por qué no?

**Figura 5.10**

22. **Pesos de aspirinas.** La figura 5.11 muestra un histograma para los pesos de 30 tabletas de aspirina elegidas al azar. La media es 665.4 miligramos. ¿Esta distribución es cercana a la normal? ¿Esta variable debe tener una distribución normal? ¿Por qué sí o por qué no?

**Figura 5.11**

23. **Áreas y frecuencias relativas.** Considere la curva normal en la figura 5.12, la cual proporciona las frecuencias relativas en una distribución de estaturas de hombres. La distri-

bución tiene una media de 69.6 pulgadas y una desviación estándar de 2.8 pulgadas.

- ¿Cuál es el área total bajo la curva?
- Estime (usando el área) la frecuencia relativa de los valores menores que 67.
- Estime la frecuencia relativa de valores mayores que 67.
- Estime la frecuencia relativa de valores entre 67 y 70.
- Estime la frecuencia relativa de valores mayores que 70.

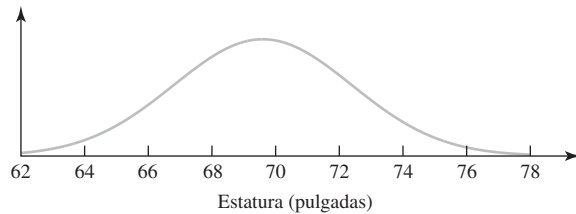


Figura 5.12

- 24. Áreas y frecuencias relativas.** Considere la curva normal en la figura 5.13, la cual proporciona las frecuencias relativas en una distribución de puntuaciones del CI. La distribución tiene una media de 100 y una desviación estándar de 16.

- ¿Cuál es el área total bajo la curva?
- Estime (usando el área) la frecuencia relativa de los valores menores que 100.
- Estime la frecuencia relativa de valores mayores que 110.
- Estime la frecuencia relativa de valores menores que 110.
- Estime la frecuencia relativa de valores entre 100 y 110.

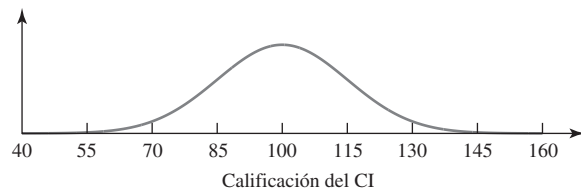


Figura 5.13

- 25. Estimación de áreas.** Considere la curva normal en la figura 5.14, la cual proporciona las frecuencias relativas en una distribución de la presión arterial sistólica para una muestra de mujeres estudiantes. La distribución tiene una desviación estándar de 14.

- ¿Cuál es la media de la distribución?
- Estime (usando el área) el porcentaje de estudiantes cuya presión sea menor que 100.

- Estime el porcentaje de estudiantes cuya presión arterial esté entre 110 y 130.
- Estime el porcentaje de estudiantes cuya presión arterial sea mayor que 130.

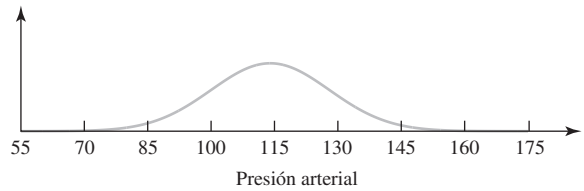


Figura 5.14

- 26. Estimación de áreas.** Considere la curva normal en la figura 5.15, la cual proporciona las frecuencias relativas en una distribución de los pesos para una muestra de estudiantes hombres.

- ¿Cuál es la media de la distribución?
- Estime (usando el área) el porcentaje de estudiantes cuyo peso sea menor que 140.
- Estime el porcentaje de estudiantes cuyo peso sea mayor que 170.
- Estime el porcentaje de estudiantes cuyo peso esté entre 140 y 160.

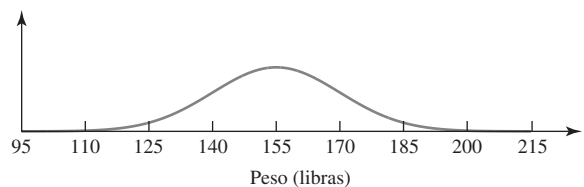


Figura 5.15



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 5 en www.aw.com/bbt.

- 27. Distribuciones de calificaciones del SAT.** El sitio web del College Board proporciona la distribución de las calificaciones del SAT (por lo regular en clases de 50 puntos). Reúna esta información y construya un histograma para cada parte del examen. Analice la validez de la afirmación que las calificaciones del SAT se distribuyen normalmente.

- 28. Determinación de distribuciones normales.** Usando las guías dadas en el texto, elija una variable que considere

debe distribuirse casi de manera normal. Reúna al menos 30 valores para la variable y construya un histograma. Comente cuánto se acerca la distribución a una distribución normal. ¿En qué difiere de una distribución normal? Trate de explicar las diferencias.

- 29. Duración de películas.** Reúna datos que apoyen o refuten la afirmación que las películas se han ido acortando a lo largo de las décadas. En específico, construya un histograma de la duración de las películas para cada década desde 1940 hasta la actualidad, determine la duración media de las películas para cada muestra y comente si estas distribuciones son normales. Analice sus resultados y proporcione razones plausibles para cualesquiera tendencias que observe.

EN LAS NOTICIAS

- 30. Distribuciones normales.** En raras ocasiones una noticia se refiere a la distribución real de una variable o indica que una variable se distribuye normalmente. Sin embargo, las variables mencionadas en las noticias deben tener alguna distribución. Encuentre dos variables en las noticias que usted sospeche tengan una distribución casi normal. Explique su razonamiento.
- 31. Distribuciones no normales.** Encuentre dos variables en las noticias que sospeche *no* tengan una distribución cercana a la normal. Explique su razonamiento.

5.2 Propiedades de la distribución normal

Considere una encuesta de la revista *Consumer Reports* en la que se preguntó a los participantes cuánto tiempo conservaban su televisor antes de reemplazarlo. La variable de interés en esta encuesta es el *tiempo de reemplazo de televisores*. Con base en la encuesta, la distribución de tiempos de reemplazo tiene una media de alrededor de 8.2 años, que denotamos como μ (la letra griega *mu*). La desviación estándar de la distribución es alrededor de 1.1 años, que denotamos como σ (la letra griega *sigma*). Haciendo la suposición razonable de que la distribución de tiempos de reemplazo de televisores es aproximadamente normal, podemos graficarla como se muestra en la figura 5.16.

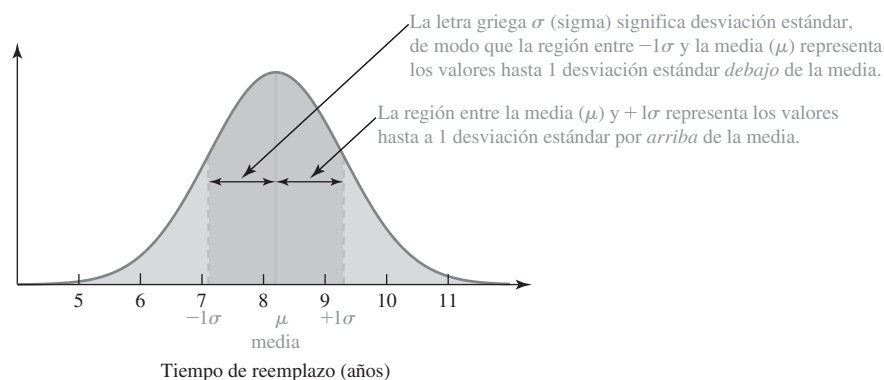


Figura 5.16 La distribución normal para tiempos de reemplazo de televisores con una media de $\mu = 8.2$ años y una desviación estándar de $\sigma = 1.1$ años.

NOTA TÉCNICA

Una distribución normal puede tener cualquier valor para la media y cualquier valor positivo para la desviación estándar. El término *distribución normal estándar* de manera específica se refiere a una distribución normal con una media de 0 y una desviación estándar de 1.

UN MOMENTO DE REFLEXIÓN

Aplique los cuatro criterios para una distribución normal (vea la sección 5.1) para explicar por qué la distribución de tiempos de reemplazo de televisores debe ser aproximadamente normal.

Puesto que todas las distribuciones normales tienen la misma forma de campana, el conocimiento de la media y de la desviación estándar nos permite decir mucho acerca de dónde se encuentran sus valores. Por ejemplo, si medimos las áreas bajo la curva en la figura 5.16, encontramos que cerca de dos tercios del área está dentro de un radio de una desviación estándar alrededor de la media, que en este caso es $8.2 - 1.1 = 7.1$ años y $8.2 + 1.1 = 9.3$ años. Por tanto, el tiempo de reemplazo de un televisor está entre 7.1 y 9.3 años para alrededor de dos tercios de las personas encuestadas. De manera similar, alrededor de 95% del área está dentro

de 2 desviaciones estándar de la media lo cual, en este caso, es entre $8.2 - 2.2 = 6.0$ años y $8.2 + 2.2 = 10.4$ años. Concluimos que el tiempo de reemplazo de televisores está entre 6.0 y 10.4 años para alrededor del 95% de la gente encuestada.

Una regla sencilla, denominada la **regla 68-95-99.7** proporciona guías precisas para el porcentaje de datos que están dentro de 1, 2 y 3 desviaciones estándar de la media para cualquier distribución normal. El recuadro siguiente indica la regla en palabras y la figura 5.17 la muestra de manera visual.

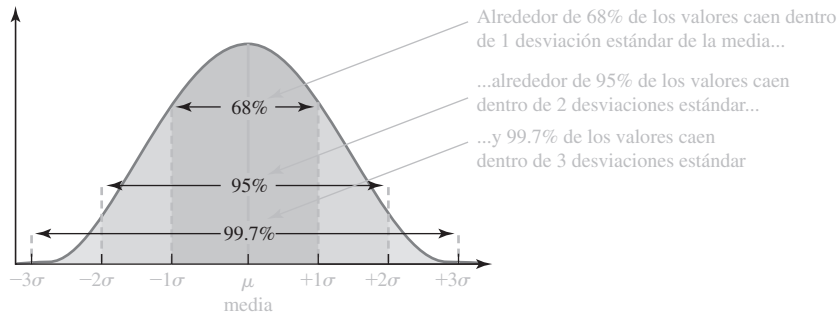


Figura 5.17 Distribución normal que ilustra la regla 68-95-99.7.

La regla 68-95-99.7 para una distribución normal

- Alrededor de 68% (o de manera más precisa, 68.3%), o un poco más de dos tercios, de los datos caen dentro de una desviación estándar de la media.
- Alrededor de 95% (o de manera más precisa, 95.4%) de los datos caen dentro de dos desviaciones estándar de la media.
- Alrededor de 99.7% de los datos caen dentro de tres desviaciones estándar de la media.

EJEMPLO 1 Calificaciones del SAT

Los exámenes del SAT que evalúan la parte verbal (lectura crítica) y la parte matemática (y el GRE, LSAT y GMAT) están diseñados de modo que sus calificaciones se distribuyan de manera normal con una media $\mu = 500$ y una desviación estándar de $\sigma = 100$. Interprete este enunciado.

Solución Con base en la regla 68-95-99.7, alrededor del 68% de los estudiantes tendrán calificaciones dentro de una desviación estándar (100 puntos) de la media de 500 puntos; esto es, alrededor de 68% de los estudiantes obtendrán calificaciones entre 400 y 600. Alrededor de 95% de las calificaciones de los estudiantes estarán dentro de 2 desviaciones estándar (200 puntos) de la media, o entre 300 y 700. Y alrededor de 99.7% de las calificaciones estarán

NOTA TÉCNICA

Como se estudia en la sección “Hablemos de psicología” al final del capítulo; la calificación media en un SAT particular puede diferir de 500 dependiendo del examen y el año que se aplica.

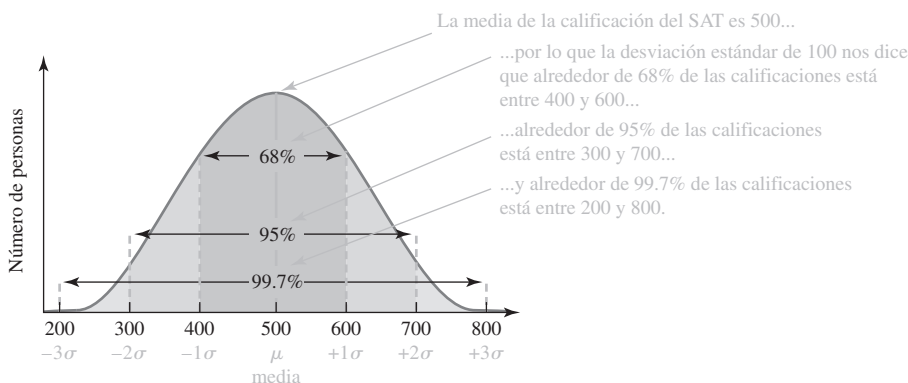


Figura 5.18 Distribución normal para calificaciones del SAT, muestra los porcentajes asociados con 1, 2 y 3 desviaciones estándar.

dentro de 3 desviaciones estándar (300 puntos) de la media, o entre 200 y 800. La figura 5.18 muestra de manera gráfica esta interpretación; observe que el eje horizontal muestra las calificaciones reales y también la distancia de la media en desviaciones estándar.

A propósito...

La Casa de Moneda de Estados Unidos emite casi mil millones de monedas de un centavo al mes. Y sin embargo, de acuerdo con un informe de la Oficina General de Contabilidad de Estados Unidos, la tasa de circulación de centavos se estima en alrededor de 34% (comparada con 88% para los cuartos de dólar). Esto significa que alrededor de dos tercios de los centavos producidos realizan un viaje en un solo sentido de la Casa de Moneda a las alcancías.



EJEMPLO 2 Detección de falsificaciones

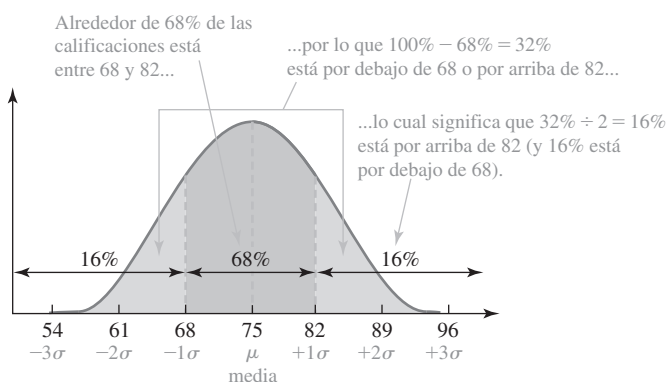
Las máquinas expendedoras pueden ajustarse para rechazar monedas por arriba y por debajo de ciertos pesos. Los pesos de monedas legales de veinticinco centavos en Estados Unidos tienen una distribución normal con media de 5.67 gramos y una desviación estándar de 0.0700 gramos. Si una máquina expendedora se ajusta para rechazar monedas que pesen más de 5.81 gramos y menos de 5.53 gramos, ¿qué porcentaje de monedas legales rechazará la máquina?

Solución Un peso de 5.81 gramos está 0.14 gramos o 2 desviaciones estándar por arriba de la media. Un peso de 5.53 gramos está 0.14 gramos, o 2 desviaciones estándar por abajo de la media. Por tanto, al sólo aceptar cuartos de dólar con peso en el rango de 5.53 a 5.81 gramos, la máquina acepta cuartos que están a dos desviaciones estándar de la media y rechaza a los que están a más de dos desviaciones estándar de la media. Mediante la regla 68-95-99.7, se tiene que 95% de los cuartos de dólar legales se aceptarán y 5% se rechazarán.

Aplicación de la regla 68-95-99.7

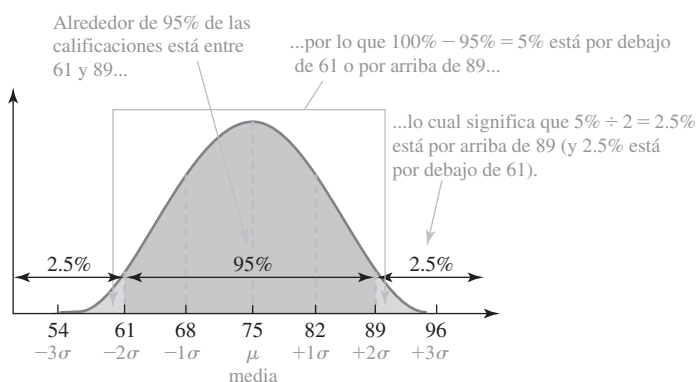
Podemos aplicar la regla 68-95-99.7 para determinar cuándo los datos caen a 1, 2 o 3 desviaciones estándar de la media. Por ejemplo, suponga que 1000 estudiantes realizan un examen y las calificaciones se distribuyen con una media de $\mu = 75$ y una desviación estándar de $\sigma = 7$.

Una calificación de 82 está 7 puntos, o una desviación estándar, por arriba de la media de 75. La regla 68-95-99.7 nos dice que alrededor de 68% de las calificaciones estarán *dentro* de una desviación estándar alrededor de la media. Se concluye que alrededor de $100\% - 68\% = 32\%$ de las calificaciones están *a más de* una desviación estándar de la media. La mitad de este 32%, o 16%, de las calificaciones está a más de una desviación estándar *por debajo* de la media; el otro 16% de las calificaciones está a más de una desviación estándar *por arriba* de la media (figura 5.19a). Así, alrededor de 16% de los 1000 estudiantes, o 160 estudiantes, obtienen más de 82.



Calificación del examen

(a)



Calificación del examen

(b)

Figura 5.19 Una distribución normal de calificaciones de un examen con media de 75 y una desviación estándar de 7. (a) 68% de las calificaciones caen dentro de 1 desviación estándar de la media. (b) 95% de las calificaciones caen dentro de 2 desviaciones estándar de la media.

De forma análoga, una calificación de 61 está 14 puntos, o 2 desviaciones estándar, por debajo de la media de 75. La regla 68-95-99.7 nos dice que alrededor de 95% de las calificaciones están *dentro* de 2 desviaciones estándar alrededor de la media, de modo que 5% de las calificaciones están *a más de* 2 desviaciones estándar de la media. La mitad de 5%, o 2.5%, de las calificaciones están a más de 2 desviaciones estándar *por abajo* de la media (figura 5.19b). Así, alrededor de 2.5% de 1000 estudiantes, o 25 estudiantes obtendrían menos de 61.

Puesto que 95% de las calificaciones caen entre 61 y 89 decimos que las calificaciones fuera de este rango no son comunes, ya que son relativamente raras.

Identificación de resultados no comunes

En estadística, con frecuencia necesitamos distinguir valores que son comunes, o “usuales”, de valores que son “raros”. Aplicando la regla 68-95-99.7, encontramos que alrededor de 95% de todos los valores de una distribución normal caen dentro de 2 desviaciones estándar de la media. Esto implica que, entre todos los valores, 5% están a más de dos desviaciones estándar de la media. Podemos utilizar esta propiedad para identificar los valores que son relativamente “raros”. **Valores raros o poco comunes** son valores que están a más de dos desviaciones estándar de la media.

EJEMPLO 3 Viaje y embarazo

De nuevo considere la pregunta de si debe avisar a una amiga embarazada que programará una importante reunión de trabajo dos semanas antes de su fecha de parto programada. Los datos reales sugieren que el número de días entre la fecha de nacimiento y la fecha programada se distribuye normalmente con una media $\mu = 0$ días, y una desviación estándar de $\sigma = 15$ días. ¿Cómo ayudaría a su amiga a tomar la decisión? ¿Un nacimiento dos semanas antes de la fecha programada se consideraría “raro”?

Solución Su amiga está suponiendo que ella *no* habrá dado a luz 14 días, o aproximadamente una desviación estándar, antes de su fecha programada. Pero como este resultado está bien dentro de 2 desviaciones estándar de la media, *no* es un resultado raro. Por la regla 68-95-99.7, el día de nacimiento para 68% de los embarazos está dentro de una desviación estándar de la media, o entre -15 y $+15$ días de la fecha programada. Esto significa que alrededor de $100\% - 68\% = 32\%$ de los nacimientos ocurrirán 15 días antes o 15 días después. Por tanto, la mitad de este 32%, o 16%, de todos los nacimientos ocurren con más de 15 días de anticipación (vea la figura 5.20). Usted debe decir a su amiga que alrededor de 16% de todos los nacimientos ocurren 15 días antes de la fecha programada. Si a su amiga le gusta pensar en términos de probabilidad, podría decirle que hay una posibilidad de 0.16 (alrededor de 1 en 6) de que ella dé a luz en o antes de la fecha de su reunión.

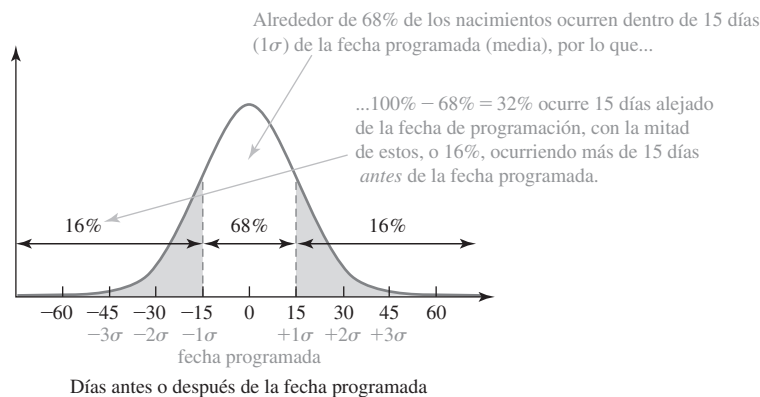


Figura 5.20 Alrededor de 16% de nacimientos ocurren con más de 15 días de anticipación de la fecha programada.

EJEMPLO 4 Ritmo cardíaco normal

Usted mide su ritmo cardíaco en reposo al mediodía de cada día durante un año y registra la información. Descubre que los datos tienen una distribución normal con una media de 66 y una desviación estándar de 4. ¿En cuántos días su ritmo cardíaco estuvo abajo de 58 pulsaciones por minuto?

Solución Un ritmo cardíaco de 58 está 8 (o 2 desviaciones estándar) por debajo de la media. De acuerdo con la regla 68-95-99.7, alrededor de 95% de los puntos están dentro de 2 desviaciones estándar de la media. Por tanto, 2.5% de los datos están a más de 2 desviaciones estándar *por debajo* de la media y 2.5% de los datos están a más de 2 desviaciones estándar *por arriba* de la media. En 2.5% de los 365 días, alrededor de 9 días, su ritmo cardíaco estaría debajo de 58 pulsaciones por minuto.

UN MOMENTO DE REFLEXIÓN

Como lo sugiere el ejemplo 4, muchas de las medidas de la prueba del ritmo cardíaco de un *solo* individuo se distribuyen normalmente. ¿Esperaría que el ritmo cardíaco promedio de *muchos* individuos se distribuya de manera normal? ¿Por qué?

EJEMPLO 5 Determinación de un percentil

Usted lleva a su hija de cuarto año de primaria al consultorio, y el médico le dice que la altura de ella es 1 desviación estándar por arriba de la media para su edad y sexo. ¿Cuál es el percentil de su estatura? Suponga que las estaturas de niñas de cuarto grado escolar se distribuyen de manera normal.

Solución Recuerde que los datos están en el percentil n de una distribución, si $n\%$ de los datos *son menores o iguales a él* (vea la sección 4.3). De acuerdo con la regla 68-95-99.7, 68% de las estaturas están dentro de una desviación estándar de la media. Por tanto, 34% de las estaturas (la mitad de 68%) están entre 0 y 1 desviación estándar *por arriba* de la media. También sabemos que, puesto que la distribución es simétrica, 50% de todas las estaturas están debajo de la media. Por tanto, $50\% + 34\% = 84\%$ de todas las estaturas son menores que 1 desviación estándar por arriba de la media (figura 5.21). Su hija está en el 84o. percentil para estaturas de niñas de cuarto grado.

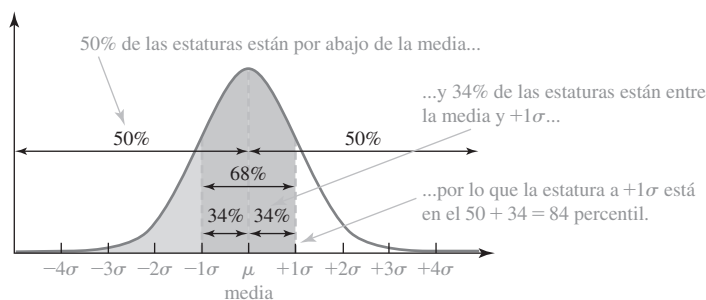


Figura 5.21 Curva de la distribución normal que muestra que 84% de las calificaciones son menores que una desviación estándar por arriba de la media.

Puntuaciones estándar

La regla 68-95-99.7 se aplica sólo a valores que están a 1, 2 o 3 desviaciones estándar de la media. Podemos generalizar esta regla si conocemos a cuántas desviaciones estándar de la media está un dato particular. El número de desviaciones estándar en las que un dato está por arriba o por debajo de la media se denomina **puntuación estándar** (o *puntuación z*), con frecuencia abreviada con la letra z . Por ejemplo,

- La puntuación estándar de la media es $z = 0$, ya que está a 0 desviaciones estándar de la media.
- La puntuación estándar de un valor a 1.5 desviaciones estándar *por arriba* de la media es $z = 1.5$.
- La puntuación estándar de un valor a 2.4 desviaciones estándar *por abajo* de la media es $z = -2.4$.

El recuadro siguiente resume el cálculo de puntuaciones estándar.

Cálculo de puntuaciones estándar

Al número de desviaciones estándar a las que un valor de datos está por arriba o por abajo de la media se le denomina su puntuación estándar (o puntuación z), y se define mediante

$$z = \text{puntuación estándar} = \frac{\text{valor} - \text{media}}{\text{desviación estándar}}$$

La puntuación estándar es positiva para valores por arriba de la media y negativa para valores menores que la media.

EJEMPLO 6 Determinación de puntuaciones estándar

El examen Stanford-Binet del CI se escala de modo que las calificaciones tengan una media de 100 y una desviación estándar de 16. Determine las puntuaciones estándar para CI de 85, 100 y 125.

Solución Calculamos las puntuaciones estándar para estos CI usando la fórmula de la puntuación estándar con una media de 100 y una desviación estándar de 16.

$$\text{puntuación estándar para 85: } z = \frac{85 - 100}{16} = -0.94$$

$$\text{puntuación estándar para 100: } z = \frac{100 - 100}{16} = 0.00$$

$$\text{puntuación estándar para 125: } z = \frac{125 - 100}{16} = 1.56$$

Podemos interpretar estas puntuaciones estándar como sigue: 85 está a 0.94 desviaciones estándar por *abajo* de la media, 100 es igual a la media y 125 está a 1.56 desviaciones estándar *por arriba* de la media. La figura 5.22 muestra estos valores en la distribución de puntuaciones de CI.

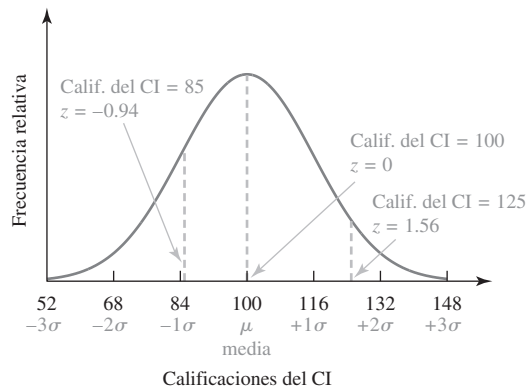


Figura 5.22 Puntuaciones estándar para calificaciones del CI de 85, 100 y 125.

Puntuaciones estándar y percentiles

Una vez que conocemos la puntuación estándar de un valor, las propiedades de la distribución normal nos permiten determinar su *percentil* en la distribución. Esto por lo común se hace con una *tabla de puntuaciones estándar*, como la tabla 5.1. (El apéndice A tiene una tabla de puntuaciones estándar más detallada). Para cada una de las puntuaciones estándar en una distribución normal, la tabla proporciona el porcentaje de valores en la distribución que son menores que el valor. Por ejemplo, la tabla muestra que 55.96% de los valores en una distribución normal tienen una puntuación estándar menor a 0.15. En otras palabras, un valor con una puntuación estándar de 0.15 está en el 55o. percentil.

EJEMPLO 7 Niveles de colesterol

Los niveles de colesterol en hombres de 18 a 24 años de edad están distribuidos normalmente con una media de 178 y una desviación estándar de 41.

- ¿Cuál es el percentil para un hombre de 20 años con un nivel de colesterol de 190?
- ¿Qué nivel de colesterol corresponde al 90o. percentil, nivel en el que puede ser necesario un tratamiento?

Solución

- La *puntuación estándar* para un nivel de colesterol de 190 es

$$z = \text{puntuación estándar} = \frac{\text{valor} - \text{media}}{\text{desviación estándar}} = \frac{190 - 178}{41} \approx 0.29$$

La tabla 5.1 muestra que una puntuación estándar de 0.29 corresponde a alrededor del 61o. percentil.

- La tabla 5.1 muestra que 90.32% de todos los datos tienen una puntuación estándar menor a 1.3. Así, el 90o. percentil está alrededor de 1.3 desviaciones estándar por arriba de la media. Dado el nivel medio de colesterol de 178 y la desviación estándar de 41, un nivel de colesterol a 1.3 desviaciones estándar de la media es

$$\underbrace{178}_{\text{media}} + \underbrace{(1.3 \times 41)}_{\substack{1.3 \text{ desviaciones} \\ \text{estándar}}} = 231.3$$

Un nivel de colesterol de alrededor de 231 corresponde al 90o. percentil.

EJEMPLO 8 Puntuaciones del CI

Las puntuaciones del CI se distribuyen de manera normal con una media de 100 y una desviación estándar de 16 (vea el ejemplo 6). ¿Cuáles son las puntuaciones del CI para personas en el percentil 75o. y el 40o. en exámenes de CI?

Solución La tabla 5.1 muestra que el percentil 75o. cae *entre* puntuaciones estándar de 0.65 y 0.70; podemos estimar que tiene una puntuación estándar de alrededor de 0.67. Esto corresponde a un CI que está 0.67 desviaciones estándar, o alrededor de $0.67 \times 16 = 11$ puntos, por arriba de la media de 100. Así, una persona en el percentil 75o. tiene un CI de 111. El percentil 40o. corresponde a una puntuación estándar de aproximadamente -0.25 , o una calificación que está $0.25 \times 16 = 4$ puntos *abajo* de la media de 100. Así, una persona en el percentil 40o. tiene un CI de 96.

EJEMPLO 9 Mujeres en el ejército

Las estaturas de las mujeres estadounidenses entre 18 y 24 años de edad se distribuyen normalmente con una media de 65 pulgadas y una desviación estándar de 2.5 pulgadas. Para alistarse en el ejército de Estados Unidos, las mujeres deben medir entre 58 y 80 pulgadas. ¿Qué porcentaje de mujeres no son elegibles para alistarse con base en sus estaturas?

Solución Las puntuaciones estándar para las estaturas mínima y máxima de 58 y 80 pulgadas son

$$\text{Para 58 pulgadas: } z = \frac{58 - 65}{2.5} = -2.8$$

$$\text{Para 80 pulgadas: } z = \frac{80 - 65}{2.5} = 6.0$$



Tabla 5.1 Puntuaciones estándar y percentiles para una distribución normal (valores de la izquierda acumulados)

Puntuación estándar	%	Puntuación estándar	%	Puntuación estándar	%	Puntuación estándar	%
-3.5	0.02	-1.0	15.87	0.0	50.00	1.1	86.43
-3.0	0.13	-0.95	17.11	0.05	51.99	1.2	88.49
-2.9	0.19	-0.90	18.41	0.10	53.98	1.3	90.32
-2.8	0.26	-0.85	19.77	0.15	55.96	1.4	91.92
-2.7	0.35	-0.80	21.19	0.20	57.93	1.5	93.32
-2.6	0.47	-0.75	22.66	0.25	59.87	1.6	94.52
-2.5	0.62	-0.70	24.20	0.30	61.79	1.7	95.54
-2.4	0.82	-0.65	25.78	0.35	63.68	1.8	96.41
-2.3	1.07	-0.60	27.43	0.40	65.54	1.9	97.13
-2.2	1.39	-0.55	29.12	0.45	67.36	2.0	97.72
-2.1	1.79	-0.50	30.85	0.50	69.15	2.1	98.21
-2.0	2.28	-0.45	32.64	0.55	70.88	2.2	98.61
-1.9	2.87	-0.40	34.46	0.60	72.57	2.3	98.93
-1.8	3.59	-0.35	36.32	0.65	74.22	2.4	99.18
-1.7	4.46	-0.30	38.21	0.70	75.80	2.5	99.38
-1.6	5.48	-0.25	40.13	0.75	77.34	2.6	99.53
-1.5	6.68	-0.20	42.07	0.80	78.81	2.7	99.65
-1.4	8.08	-0.15	44.04	0.85	80.23	2.8	99.74
-1.3	9.68	-0.10	46.02	0.90	81.59	2.9	99.81
-1.2	11.51	-0.05	48.01	0.95	82.89	3.0	99.87
-1.1	13.57	0.0	50.00	1.0	84.13	3.5	99.98

Nota: la tabla muestra los percentiles para puntuaciones estándar entre -3.5 y 3.5 , aunque son posibles puntuaciones estándar menores y mayores. (El apéndice A tiene una tabla de puntuaciones estándar más detallada). La columna % proporciona el porcentaje de valores en la distribución menores que la puntuación estándar correspondiente.

La tabla 5.1 muestra que una puntuación estándar de -2.8 corresponde al percentil 0.26. Una puntuación estándar de 6.0 no aparece en la tabla 5.1, lo que significa que está por arriba del percentil 99.980. (el percentil más alto mostrado en la tabla). Así, 0.26% de todas las mujeres serán demasiado bajas para servir en el ejército y menos de 0.02% de todas las mujeres son demasiado altas para servir en el ejército. Juntas, menos de 0.28% de todas las mujeres, o alrededor de 1 de cada 400 mujeres no son elegibles para servir en el ejército, con base en sus estaturas.

Hacia la probabilidad

Suponga que selecciona un bebé al azar y pregunta si el bebé nació con más de 15 días antes de su fecha programada. Puesto que los nacimientos se distribuyen de manera normal alrededor de la fecha programada con una desviación estándar de 15 días, sabemos que 16% de todos los nacimientos ocurren más de 15 días antes de la fecha programada (vea el ejemplo 3). Por tanto, para un bebé elegido al azar, podemos decir que existen 0.16 posibilidades (alrededor de 1 en 6) que el bebé haya nacido con más de 15 días de anticipación. En otras palabras, las propiedades de la distribución normal nos permiten hacer una *afirmación de probabilidad* acerca de un individuo. En este caso, nuestra afirmación es que la probabilidad de que un nacimiento ocurra con mayor antelación a 15 días es 0.16.

El ejemplo muestra que las propiedades de la distribución normal pueden volver a enunciarse en términos de ideas de probabilidad. Por esta razón dedicaremos el capítulo siguiente a estudiar las ideas fundamentales de la probabilidad. Pero primero, en la sección siguiente, utilizaremos las ideas básicas que hemos estudiado hasta ahora para introducir uno de los conceptos más importantes en estadística.

Sección 5.2 Ejercicios

Alfabetización estadística y pensamiento crítico

1. **Puntuación estándar.** ¿Qué es la puntuación z para la media en un conjunto de valores que se distribuyen normalmente?
2. **Puntuación estándar.** El porcentaje total de valores a la izquierda de una puntuación z particular es 75%. ¿Cuál es el porcentaje de valores a la derecha de la misma puntuación z ?
3. **Distribuciones.** Para el tiro de un dado, el resultado medio es 3.5. ¿Podemos aplicar la regla 68-95-99.7 y concluir que 95% de todos los resultados caen dentro de 2 desviaciones estándar de 3.5? ¿Por qué sí o por qué no?
4. **Puntuaciones z y porcentajes.** La tabla 5.1 incluye puntuaciones estándar y percentiles. ¿Una puntuación z puede tener un valor negativo? ¿Un percentil puede tener un valor negativo? Explique.

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Prueba psicológica.** Todas las puntuaciones en una prueba psicológica son positivas, se distribuyen de manera normal con una media de 55 y una desviación estándar de 65.
6. **Pesos al nacer.** Los pesos al nacer en Estados Unidos se distribuyen normalmente con una media de 3 420 gramos y una desviación estándar de 495 gramos.
7. **Calificaciones del SAT.** Las calificaciones del SAT se distribuyen normalmente con una media de 1518 y una desviación estándar de 325.
8. **Calificaciones del SAT.** Las calificaciones del SAT se distribuyen normalmente con una media de 1518 y una desviación estándar de 325, y la calificación de Marc de 1 840 es inusualmente alta.

Conceptos y aplicaciones

9. **Uso de la regla 68-95-99.7.** Suponga que un conjunto de calificaciones se distribuye normalmente con una media de 100 y una desviación estándar de 20. Utilice la regla 68-95-99.7 para determinar las cantidades siguientes.
 - a. Porcentaje de calificaciones menores que 100.
 - b. Frecuencia relativa de calificaciones menores que 120.
 - c. Porcentaje de calificaciones menores que 140.
 - d. Porcentaje de calificaciones menores que 80.
 - e. Frecuencia relativa de calificaciones menores que 60.

- f. Porcentaje de calificaciones mayores que 120.
- g. Porcentaje de calificaciones mayores que 140.
- h. Frecuencia relativa de calificaciones mayores que 80.
- i. Porcentaje de calificaciones entre 80 y 120.
- j. Porcentaje de calificaciones entre 80 y 140.

10. **Uso de la regla 68-95-99.7.** Suponga que el ritmo cardíaco en reposo para una muestra de individuos se distribuye normalmente con una media de 70 y una desviación estándar de 15. Utilice la regla 68-95-99.7 para determinar las cantidades siguientes.

- a. El porcentaje de ritmos menores que 70.
- b. El porcentaje de ritmos menores que 55.
- c. La frecuencia relativa de ritmos menores que 40.
- d. El porcentaje de ritmos menores que 85.
- e. La frecuencia relativa de ritmos menores que 100.
- f. El porcentaje de ritmos mayores que 85.
- g. El porcentaje de ritmos mayores que 55.
- h. La frecuencia relativa de ritmos mayores que 40.
- i. El porcentaje de ritmos entre 55 y 85.
- j. El porcentaje de ritmos entre 70 y 100.

11. **Aplicación de la regla 68-95-99.7.** En un estudio de comportamiento facial, a personas en un grupo de control se les tomó el tiempo para que hicieran contacto visual en un periodo de 5 minutos. Sus tiempos se distribuyeron normalmente con una media de 184.0 segundos y una desviación estándar de 55.0 segundos (con base en información de “Ethological Study of Facial Behavior in Nonparanoid and Paranoid Schizophrenic Patients”, de Pittman, Olk Orr y Singh. *Psychiatry*, volumen 144, número 1). Utilice la regla de 68-95-99.7 para determinar la cantidad indicada.

- a. Determine el porcentaje de tiempos dentro de 55.0 segundos de la media de 184.0 segundos.
- b. Determine el porcentaje de tiempos dentro de 110.0 segundos de la media de 184.0 segundos.
- c. Determine el porcentaje de tiempos dentro de 165.0 segundos de la media de 184.0 segundos.
- d. Determine el porcentaje de tiempos entre 184.0 y 294 segundos.

12. **Aplicación de la regla 68-95-99.7.** Cuando se diseña la ubicación de un reproductor de CD en un modelo nuevo de automóvil, los ingenieros deben considerar el alcance del conductor al tablero. Las mujeres tienen un alcance que se distribuye normalmente con una media de 27.0 pulgadas y

una desviación estándar de 1.3 pulgadas (con base en información de una encuesta antropométrica de Gordon, Churchill, *et al*). Utilice la regla 68-95-99.7 para determinar la cantidad indicada.

- Determine el porcentaje de mujeres con alcance al tablero entre 24.4 y 29.6 pulgadas.
- Determine el porcentaje de mujeres con alcance menor que 30.9 pulgadas.
- Determine el porcentaje de mujeres con alcance al tablero entre 27.0 y 28.3 pulgadas.

Calificaciones del CI. Para los ejercicios del 13 al 24 utilice la distribución normal de calificaciones de CI, que tiene una media de 100 y una desviación estándar de 16. Utilice la tabla 5.1 (en la página 211) para determinar las cantidades indicadas. Nota: la tabla 5.1 muestra puntuaciones z de -3.5 a $+3.5$. En estos problemas, para puntuaciones estándar superiores a 3.5 utilice un percentil de 99.99%; para todas las puntuaciones inferiores a -3.5 utilice un percentil de 0.01%.

- Porcentaje de calificaciones menores que 100.
- Porcentaje de calificaciones menores que 84.
- Porcentaje de calificaciones mayores que 116.
- Porcentaje de calificaciones menores que 76.
- Porcentaje de calificaciones mayores que 132.
- Porcentaje de calificaciones menores que 80.
- Porcentaje de calificaciones menores que 65.
- Porcentaje de calificaciones menores que 129.
- Porcentaje de calificaciones entre 84 y 116.
- Porcentaje de calificaciones entre 76 y 124.
- Porcentaje de calificaciones entre 80 y 120.
- Porcentaje de calificaciones entre 92 y 115.

Estaturas de mujeres. Para los ejercicios del 25 al 36 utilice la distribución normal de estaturas de mujeres adultas, la cual tiene media de 162 centímetros y una desviación estándar de 6 centímetros. Utilice la tabla 5.1 para determinar las cantidades indicadas. Nota: la tabla 5.1 muestra puntuaciones z de -3.5 a $+3.5$. En estos problemas, para puntuaciones estándar superiores a 3.5, utilice un percentil de 99.99%; para todas las puntuaciones inferiores a -3.5 , utilice un percentil de 0.01%.

- El porcentaje de estaturas mayores que 162 centímetros.
- El porcentaje de estaturas menores que 168 centímetros.
- El porcentaje de estaturas mayores que 156 centímetros.
- El porcentaje de estaturas mayores que 171 centímetros.
- El porcentaje de estaturas menores que 177 centímetros.
- El porcentaje de estaturas menores que 147 centímetros.
- El porcentaje de estaturas menores que 144 centímetros.

- El porcentaje de estaturas mayores que 179 centímetros.
- El porcentaje de estaturas entre 156 y 168 centímetros.
- El porcentaje de estaturas entre 159 y 165 centímetros.
- El porcentaje de estaturas entre 148 y 170 centímetros.
- El porcentaje de estaturas entre 146 y 156 centímetros.
- Pesos de monedas.** Considere la tabla siguiente, que muestra la media oficial del peso y la desviación estándar estimada para las cinco monedas de Estados Unidos. Suponga que una máquina expendedora está diseñada para rechazar todas las monedas con pesos que estén a más de dos desviaciones estándar, por arriba o por abajo, de la media. Para cada moneda determine el rango de pesos que son aceptables para las máquinas expendedoras. En cada caso, ¿qué porcentaje de monedas legales son rechazadas por la máquina?

Moneda	Peso (gramos)	Desviación estándar estimada (gramos)
Centavo (penique)	2.500	0.03
Cinco centavos (níquel)	5.000	0.06
Diez centavos (dime)	2.268	0.03
25 centavos (cuarto de dólar)	5.670	0.07
50 centavos (medio dólar)	11.340	0.14

- Duración de embarazos.** La duración de los embarazos se distribuyen de manera normal con una media de 268 días y una desviación estándar de 15 días.
 - ¿Cuál es el porcentaje de embarazos que dura menos de 250 días?
 - ¿Cuál es el porcentaje de embarazos que dura más de 300 días?
 - Un nacimiento se considera prematuro si el embarazo dura menos de 238 días, ¿cuál es el porcentaje de nacimientos prematuros?
- Calificaciones del SAT.** Con base en información del College Board, las calificaciones se distribuyen normalmente con una media de 1518 y una desviación estándar de 325.
 - Determine el porcentaje de calificaciones del SAT mayores que 2000.
 - Determine el porcentaje de calificaciones del SAT menores que 1500.
 - Determine el porcentaje de calificaciones del SAT entre 1600 y 2100.
- Calificaciones del GRE.** Suponga que las calificaciones en el examen de registro de graduados (GRE, por sus siglas en inglés) se distribuyen normalmente con una media de 497 y una desviación estándar de 115.

- a. Para ser admitido, una escuela de graduados pide una calificación GRE de 650. ¿A qué percentil corresponde?
 - b. Para ser admitido, una escuela de graduados pide una calificación GRE en el percentil 95o. ¿A qué calificación real corresponde?
- 41. Calibración de barómetros.** Los investigadores para un productor de barómetros (dispositivos para medir la presión atmosférica) leen cada uno de 50 barómetros en el mismo momento del día. La media de las lecturas es 30.4 (pulgadas de mercurio) con una desviación estándar de 0.23 pulgadas y las lecturas parece que se distribuyen normalmente.
- a. ¿Qué porcentaje de la lectura de los barómetros son mayores que 31?
 - b. ¿Qué porcentaje de la lectura de los barómetros son menores que 30?
 - c. La compañía decide rechazar los barómetros cuya lectura esté a más de 1.5 desviaciones estándar, por arriba o por abajo, de la media. ¿Cuál es la lectura crítica inferior a la que los barómetros son rechazados? ¿Cuál es la lectura crítica superior a la que los barómetros son rechazados?
 - d. ¿Cuál lectura tomaría como la presión atmosférica real en el momento en el que se tomó la lectura de los barómetros? Explique.
- 42. Calificaciones en prueba de ortografía.** En el examen distrital de ortografía, las 60 niñas tienen una calificación media de 71 puntos con una desviación estándar de 6, mientras que los 50 niños tienen una calificación media de 66 puntos con una desviación estándar de 5 puntos. Aquellos estudiantes con una calificación mayor que 75 son elegibles para ir a la fase estatal del concurso. ¿Qué porcentaje de aquellos que van a la fase estatal son niñas?
- 43. Convirtiéndose en infante de marina.** De acuerdo con datos de la Encuesta Nacional de Salud, las estaturas de hombres adultos se distribuyen normalmente con una media de 69.0 pulgadas y una desviación estándar de 2.8 pulgadas. Los cuerpos de marina de Estados Unidos requieren que los hombres tengan una estatura entre 64 y 78 pulgadas. ¿Qué porcentaje de estadounidenses son elegibles para ser infantes de marina, con base en la estatura?
- 44. Duración de películas.** Con base en una muestra aleatoria de duraciones de películas, la duración media es de 110.5 minutos con una desviación estándar de 22.4 minutos. Suponga que la duración de las películas se distribuye normalmente.

- a. ¿Qué fracción de las películas tienen una duración mayor a 2 horas?
- b. ¿Qué fracción de las películas duran menos de una hora y media?
- c. ¿Cuál es la probabilidad de que una película elegida aleatoriamente tendrá una duración de menos de dos horas y media?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 5 en www.aw.com/bbt.

- 45. Biblioteca de Datos e Historia.** Visite el sitio web de Data and Story Library y encuentre un ejemplo (o muchos) de una distribución normal. Escriba una página acerca de sus hallazgos. Incluya una descripción cuidadosa de la variable bajo consideración, detalles de la distribución y una explicación de por qué espera que esta variable tenga aproximadamente una distribución normal.
- 46. Estadísticas vitales por estado.** El Centro Nacional para Estadísticas de Salud de Estados Unidos proporciona estadísticas vitales por estado (y condado) bajo el rubro State Tabulated Data. Encuentre una variable (por ejemplo natalidad, mortalidad, o mortalidad infantil). Construya un histograma que muestre esta variable para cada estado. Comente si la variable está distribuida normalmente por estado.
- 47. Demostraciones de la distribución normal en la web.** Haga una búsqueda en internet sobre las palabras clave “distribución normal” y encuentre una demostración animada de la distribución normal. Describa cómo funciona la demostración y las características útiles que observó.
- 48. Estimación de un minuto.** Pida a sujetos de estudio que estimen un minuto sin mirar un reloj o un cronómetro. Cada sujeto debe decir “vamos” al inicio del minuto y luego “alto” cuando consideren que ha transcurrido un minuto. (De manera alterna, podría usted repetidamente hacer el cronometraje observando cualquier reloj o cronómetro y luego viendo el tiempo correcto cuando considere que ha transcurrido un minuto). Utilice un reloj para registrar los tiempos reales. Construya una gráfica de las estimaciones. ¿La gráfica parece aproximadamente normal? ¿La media parece que es cercana a un minuto?

5.3 El teorema del límite central

Una maestra de preparatoria tiene 100 alumnos de último año que presentan un examen de colocación para la universidad. El examen está diseñado para que tenga una calificación media de 500 y una desviación estándar de 100. Con los métodos de este capítulo ella puede determinar el porcentaje de calificaciones individuales que están probablemente por arriba, digamos, 600. Pero, ¿ella puede pronosticar algo acerca del desempeño de su grupo de 100 estudiantes? Por ejemplo, ¿cuál es la probabilidad de que la calificación media del grupo estará por arriba de 600? Este tipo de preguntas, en la que cuestionamos acerca de la calificación media de un grupo o muestra sacada de una población mucho mayor, puede responderse con el *teorema del límite central*.

Antes de que obtengamos el teorema mismo, podemos desarrollar una comprensión considerando el tiro de dados. Suponga que tira *un* dado 1 000 veces y registra el resultado de cada tiro, que puede ser el número 1, 2, 3, 4, 5 o 6. La figura 5.23 muestra un histograma de resultados. Los seis resultados tienen casi la misma frecuencia relativa, ya que es igualmente probable que el dado caiga en cada uno de sus seis posibles formas. Esto es, el histograma muestra una *distribución* (casi) *uniforme* (vea la sección 4.2). Con los métodos descritos en el capítulo 4 podemos calcular la media y la desviación estándar para esta distribución. Resulta que la distribución en la figura 5.23 tiene una media de 3.41 y una desviación estándar de 1.73.

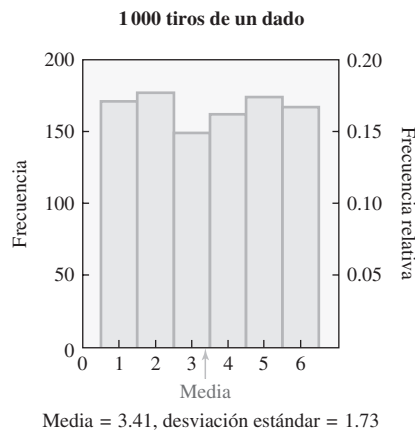


Figura 5.23 Distribución de la frecuencia y de la frecuencia relativa de los resultados del tiro de un dado 1 000 veces.

Ahora suponga que tiramos *dos* dados 1 000 veces y registramos la *media* de los dos números que aparezcan en cada tiro (figura 5.24). Para determinar la media de un solo tiro sumamos los dos números y dividimos entre 2. Por ejemplo, si los dos dados salen 3 y 5, la media del tiro es $(3 + 5)/2 = 4$. Los valores posibles de la media en un tiro de dos dados son 1.0, 1.5, 2.0, ..., 5.0, 5.5, 6.0.

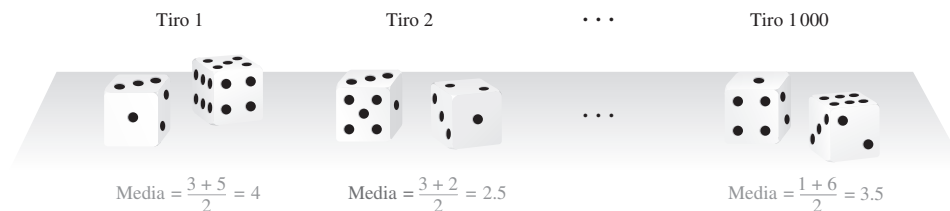


Figura 5.24 Este diagrama representa la idea de tirar dos dados 1 000 veces y registrar la media de cada tiro. La media de los valores de dos dados es su suma dividida entre 2. Esta media puede variar de $(1 + 1)/2 = 1$ a $(6 + 6)/2 = 6$.

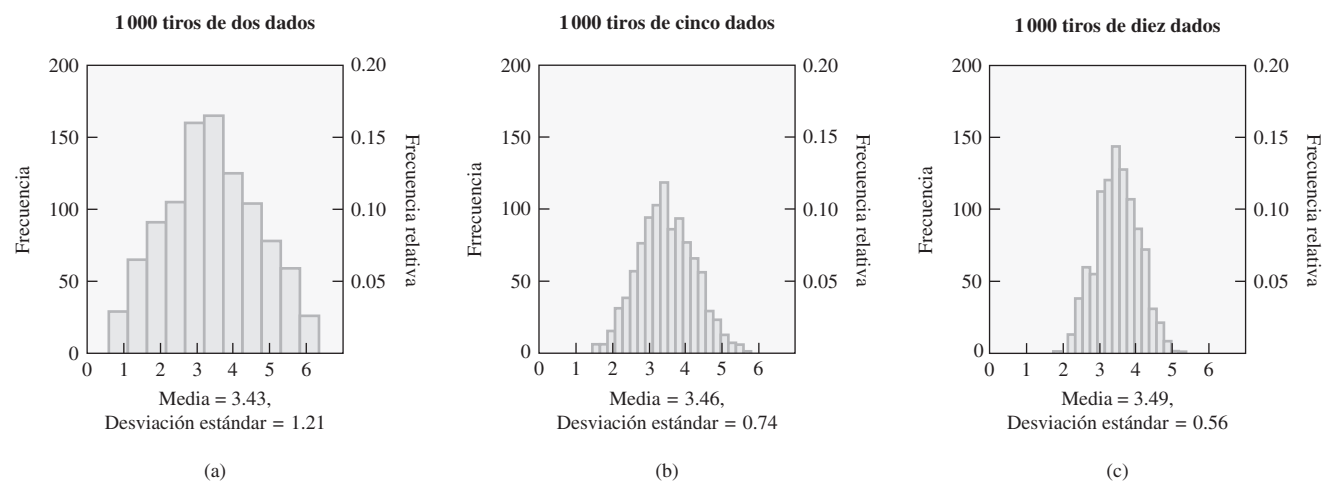


Figura 5.25 Las distribuciones de frecuencias y frecuencias relativas de las medias muestrales del tiro de (a) dos dados 1000 veces, (b) cinco dados 1 000 veces, y (c) diez dados 1 000 veces.

La figura 5.25a muestra el resultado de tirar dos dados 1000 veces. Los valores más comunes en esta distribución son los valores centrales 3.0, 3.5 y 4.0. Estos valores son comunes porque pueden ocurrir de varias maneras. Por ejemplo, una media de 3.5 ocurre si los dados caen como 1 y 6, 2 y 5, 3 y 4, 4 y 3, 5 y 2 o bien 6 y 1. Valores altos y bajos ocurren con menos frecuencia ya que pueden ocurrir de pocas formas. Por ejemplo, un tiro puede tener una media de 1.0 sólo si ambos dados caen mostrando un 1. De nuevo, es posible calcular la media y la desviación estándar para esta distribución, y ellas resultan ser 3.43 y 1.21, respectivamente.

¿Qué sucede si aumentamos el número de dados lanzados? Suponga que tiramos cinco dados 1000 veces y registramos la media de los cinco números en cada tiro. Un histograma para este experimento se muestra en la figura 5.25b. Una vez más, los valores centrales alrededor de 3.5 ocurren con más frecuencia pero la dispersión de la distribución es más reducida que en los dos casos anteriores. Al calcular la media y la desviación estándar de esta distribución obtenemos los valores de 3.46 y 0.74, respectivamente.

Si aumentamos el número de dados a diez en cada uno de los 1000 tiros, encontramos el histograma de la figura 5.25c, que es aún más angosto. En este caso, la media es 3.49 y la desviación estándar es 0.56.

La tabla 5.2 resume los cuatro experimentos descritos. Las columnas 2 y 3 de la tabla se refieren a una *distribución de medias*, ya que en cada uno de los experimentos de tiro de dados registramos la media de cada uno de los 1000 tiros (esto es, la media de un dado en cada tiro, de dos dados en cada tiro, de cinco dados en cada tiro y de diez dados en cada tiro). Así, la media para los 1000 tiros en un experimento es una *media de la distribución de las medias* (columna 2). De manera similar, la desviación estándar para los 1000 tiros en un experimento es una *desviación estándar de la distribución de las medias* (columna 3).

Tabla 5.2 Resumen de los experimentos de tiro de dados		
Número de dados lanzados cada vez	Media de la distribución de las medias	Desviación estándar de la distribución de las medias
1	3.41	1.73
2	3.43	1.21
5	3.46	0.74
10	3.49	0.56

Una idea sorprendente surge de estos cuatro experimentos. Tirar $n = 1$ dado 1 000 veces puede considerarse como tomar 1 000 muestras de tamaño $n = 1$ de la población de todos los posibles tiros de dados. Tirar $n = 2$ dados 1 000 veces puede verse como tomar 1 000 muestras de tamaño $n = 2$. Asimismo, tirar $n = 5$ y $n = 10$ dados 1 000 veces es como tomar 1 000 muestras de tamaño $n = 5$ y $n = 10$, respectivamente. La tabla 5.2 muestra que cuando el tamaño de la muestra aumenta, la media de la distribución de medias se aproxima al valor 3.5 y la desviación estándar se hace más pequeña (haciendo que la distribución sea más angosta). Más importante aún, la distribución se hace cada vez más parecida a la distribución normal conforme el tamaño de la muestra aumenta. Este último hecho puede parecer sorprendente, ya que hemos tomado muestras de una distribución *uniforme* (los resultados del tiro de un dado mostrados en la figura 5.23), *no* de una distribución normal. Sin embargo, la distribución de las medias claramente se aproxima a una distribución normal para tamaños grandes de la muestra. Este hecho es una consecuencia del **teorema del límite central**.

Teorema del límite central

Suponga que tomamos muchas muestras aleatorias de tamaño n para una variable con cualquier distribución (no necesariamente una distribución normal) y registra la distribución de las *medias* de cada muestra. Entonces,

1. La distribución de las medias será aproximadamente una distribución normal para muestras de tamaño grande.
2. La media de la distribución de las medias se aproxima a la media de la población, μ , para muestras de tamaño grande.
3. La desviación estándar de la distribución de las medias se aproxima a $\sigma/\sqrt{n}$ para muestras de tamaño grande, donde σ es la desviación estándar de la población.

Asegúrese de observar el ajuste muy importante, descrito en el punto 3 anterior, que debe hacerse cuando se trabaja con muestras o grupos en lugar de individuos:

La desviación estándar de la distribución de medias muestrales no es la desviación estándar de la población, σ , en su lugar $\sigma/\sqrt{n}$, donde n es el tamaño de las muestras.

NOTA TÉCNICA

(1) Para propósitos prácticos, la distribución de las medias será casi normal si el tamaño de la muestra es mayor que 30. (2) Si la población original se distribuye normalmente, entonces las medias muestrales serán normalmente distribuidas para *cualquier* tamaño de muestra n . (3) En el caso ideal, cuando la distribución de medias esté formada de *todas* las muestras posibles, la media de la distribución de medias es *igual* a μ y la desviación estándar de la distribución de medias es *igual* a $\sigma/\sqrt{n}$.

UN MOMENTO DE REFLEXIÓN

Confirme que las desviaciones estándar de las distribuciones de las medias dadas en la tabla 5.2 para $n = 2, 5, 10$ coinciden con el pronóstico del teorema del límite central, dado que $\sigma = 1.73$ (la desviación estándar de la población encontrada en la figura 5.23). Por ejemplo, con $n = 2$, $\sigma/\sqrt{2} = 1.22 \approx 1.21$.

Resumamos los ingredientes del teorema del límite central. Siempre iniciamos con una variable particular, tal como los resultados del tiro de unos dados o los pesos de personas, que varían aleatoriamente en una población. La variable tiene cierta media, μ , y desviación estándar, σ , que podemos o no conocer. Esta variable tiene *cualquier* clase de distribución, no necesariamente normal. Ahora tomamos muchas muestras de esa variable, con n elementos en cada muestra, y encontramos la media de cada muestra (tal como el valor medio de n dados o el peso medio de una muestra de n personas). Si luego hacemos un histograma de las medias de todas las muestras, veremos una distribución que es cercana a una distribución normal. Entre mayor sea el tamaño de la muestra, n , más cercana será la distribución de las medias a una distribución normal. Un estudio cuidadoso de la figura 5.26 debe ayudar a consolidar estas ideas importantes.

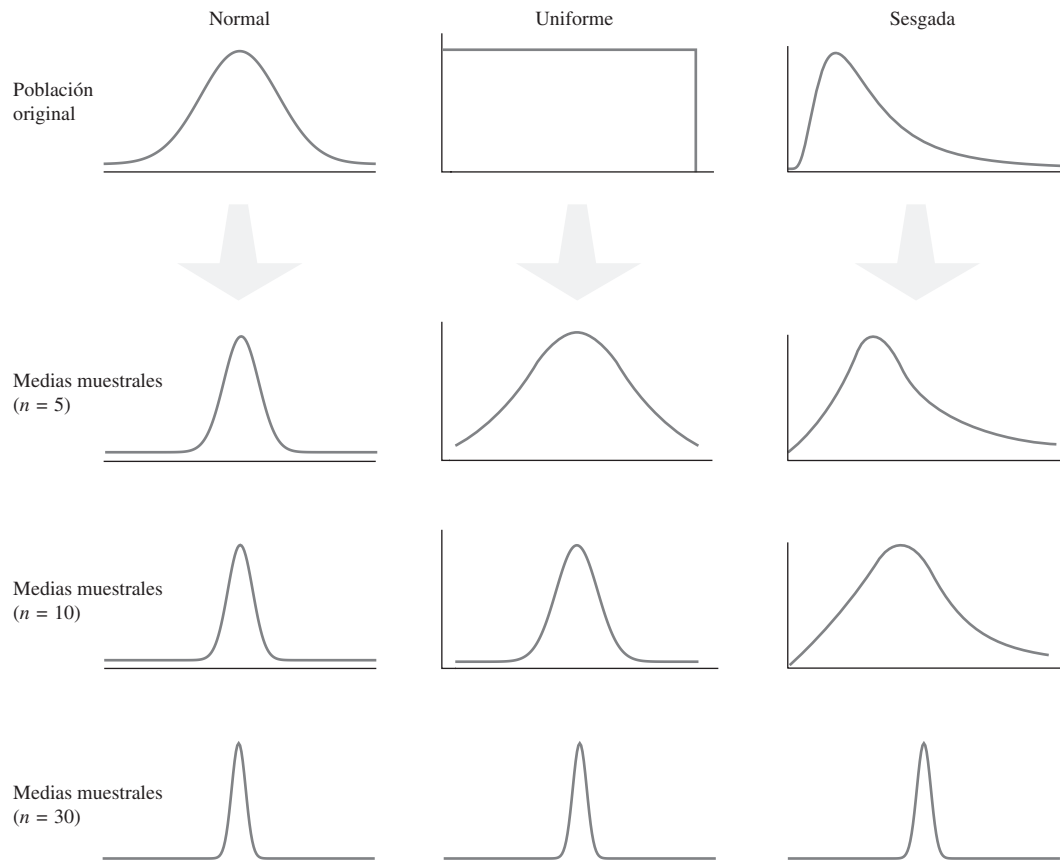


Figura 5.26 Cuando el tamaño de la muestra crece ($n = 5, 10, 30$), la distribución de las medias muestrales tiende a una distribución normal, sin importar la forma de la distribución original. Entre más grande sea el tamaño de la muestra, más pequeña es la desviación estándar de la distribución de las medias muestrales.

EJEMPLO 1 Pronóstico de calificaciones en exámenes

Usted es el director de una escuela secundaria y sus 100 alumnos de octavo grado están a punto de presentar un examen nacional estandarizado. Este examen está diseñado para que la calificación media sea $\mu = 400$ con una desviación estándar de $\sigma = 70$. Suponga que las calificaciones se distribuyen normalmente.

- ¿Cuál es la posibilidad de que *uno* de sus alumnos de octavo grado, seleccionado al azar, tenga una calificación por debajo de 375 en el examen?
- Su desempeño como director depende de las buenas calificaciones obtenga que todo su *grupo* en el examen. ¿Cuál es la posibilidad de que su grupo de 100 alumnos de octavo grado tenga una calificación *media* por debajo de 375?

Solución

- Al tratar con una calificación individual, utilizamos el método de las puntuaciones estándar, estudiado en la sección 5.2. Dada la media de 400 y desviación estándar de 70, una calificación de 375 tiene una puntuación estándar de

$$z = \frac{\text{valor} - \text{media}}{\text{desviación estándar}} = \frac{375 - 400}{70} = -0.36$$

De acuerdo con la tabla 5.1, una puntuación estándar de -0.36 corresponde a alrededor del 36.º percentil, esto es, 36% de todos los estudiantes puede esperarse que obtengan menos de 375. Por tanto, existe una posibilidad de 0.36 de que un alumno seleccionado al azar tenga

una calificación inferior a 375. Observe que necesitamos saber que las calificaciones tienen una distribución normal para hacer este cálculo, ya que la tabla de puntuaciones estándar sólo se aplica a las distribuciones normales.

- b. La pregunta acerca de la media de un *grupo* de estudiantes debe manejarse con el teorema del límite central. De acuerdo con este teorema, si tomamos muestras aleatorias de tamaño $n = 100$ estudiantes y calculamos la calificación media de los exámenes de cada grupo, la distribución de medias es aproximadamente normal. Además, la media de esta distribución es $\mu = 400$ y su desviación estándar es $\sigma/\sqrt{n} = 70/\sqrt{100} = 7$. Con estos valores para la media y la desviación estándar, la puntuación estándar para una calificación media en el examen de 375 es

$$z = \frac{\text{valor} - \text{media}}{\text{desviación estándar}} = \frac{375 - 400}{7} = -3.57$$

La tabla 5.1 muestra que una puntuación estándar de -3.5 corresponde al percentil 0.02, y la puntuación estándar en este caso es aún menor. En otras palabras, menos de 0.02% de todas las muestras aleatorias de 100 estudiantes tendrán una calificación media menor a 375. Por tanto, la posibilidad de que un grupo, de 100 estudiantes, seleccionado al azar tenga una calificación media inferior a 375 es menor a 0.0002, o alrededor de 1 en 5000. Observe que este cálculo que considera la media del grupo *no* depende de que las calificaciones individuales tengan una distribución normal.

Este ejemplo tiene una lección importante. La posibilidad de una calificación *individual* abajo de 375 es más de 1 en 3 (36%), pero la posibilidad de que un *grupo* de 100 estudiantes tenga una calificación media inferior a 375 es menor de 1 en 5000 (0.02%). En otras palabras, hay mucha más variación en las calificaciones individuales que en las medias de grupos de individuos.

EJEMPLO 2 Igualdad en salarios

El salario medio de los 9000 empleados en Holley.com es $\mu = \$26400$ con una desviación estándar de $\sigma = \$2420$. Un encuestador muestrea a 400 empleados seleccionados aleatoriamente y determina que el salario medio de la muestra es \$26650. ¿Es probable que el encuestador obtuviese estos resultados por el azar, o la discrepancia sugiere que los resultados del encuestador son sospechosos?

Solución La pregunta trata con la media de un *grupo* de 400 individuos, que es un caso para el teorema del límite central. El teorema nos dice que si seleccionamos muchos grupos de 400 individuos y calculamos la media de cada grupo, la distribución de las medias será cercana a una normal con media de $\mu = \$26400$ y una desviación estándar de $\sigma/\sqrt{n} = \$2420/\sqrt{400} = \121 . Dentro de la distribución de medias, un salario medio de \$26650 tiene una *puntuación estándar* de

$$z = \frac{\text{valor} - \text{media}}{\text{desviación estándar}} = \frac{\$26650 - \$26400}{\$121} = 2.07$$

En otras palabras, si suponemos que la muestra se seleccionó de manera aleatoria, su salario medio está a más de dos desviaciones estándar por arriba del salario medio del de toda la compañía. De acuerdo con la tabla 5.1, una puntuación estándar de 2.07 cae cerca del 98o. percentil. Así, el salario medio de *esta* muestra es mayor que el salario medio que encontraríamos en 98% de las posibles muestras de 400 trabajadores. Esto es, la posibilidad de seleccionar un grupo de 400 trabajadores con un salario medio por arriba de \$26650 es alrededor de 2%, o 0.02. El salario medio de la muestra es sorprendentemente alto; tal vez la encuesta tuvo fallas.

UN MOMENTO DE REFLEXIÓN

¿Un salario de \$26650 para un *individuo* estará arriba o abajo del percentil 98o.? Explique.

El valor del teorema del límite central

El teorema del límite central nos permite decir algo acerca de la media de un grupo si conocemos la media, μ , la desviación estándar, σ , de la población completa. Esto puede ser útil, pero resulta que la aplicación opuesta es mucho más importante.

Dos actividades principales de la estadística son realizar estimaciones de las medias de la población y probar afirmaciones acerca de las medias de la población. Suponga que usted *no* conoce la media de una variable para toda la población. ¿Es posible hacer una buena estimación de la media de la población (tal como el ingreso medio de todos los usuarios de internet) si sólo se conoce la media de una muestra mucho más pequeña? Como probablemente lo ha adivinado, en ser capaces de responder este tipo de preguntas radica el corazón del muestreo estadístico, en especial en encuestas y sondeos de opinión. El teorema del límite central proporciona la clave para responder tales preguntas. Regresaremos a este tema en el capítulo 8.

Sección 5.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Muestreo de conveniencia.** Puesto que una estudiante de estadística esperó hasta el último momento para hacer su proyecto, sólo tiene tiempo suficiente para recolectar estaturas de amigas y parientes mujeres. Luego calcula la estatura media de las mujeres en su muestra. Suponiendo que las mujeres tienen estaturas que se distribuyen normalmente con una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas, ¿puede utilizar el teorema del límite central cuando analiza la estatura media de su muestra?
- Notación.** En esta sección observó que la desviación estándar de las medias muestrales es $\sigma/\sqrt{n}$. En esa expresión, ¿qué representa σ y qué representa n ?
- Teorema del límite central.** Un proceso consiste en la repetición de la operación siguiente: seleccionar aleatoriamente dos valores de alguna población y luego encontrar la media de los dos valores. ¿El teorema del límite central sugiere que las medias muestrales se distribuirán normalmente? ¿Por qué sí o por qué no?
- Teorema del límite central.** Un proceso consiste en la repetición de esta operación: seleccionar aleatoriamente cuatro valores de una población con una distribución normal y una desviación estándar σ . ¿Cuál es la desviación estándar de las medias muestrales?

Conceptos y aplicaciones

- Puntuaciones del CI y el teorema del límite central.** Las calificaciones del CI se distribuyen normalmente con una media de 100 y una desviación estándar de 16. Suponga que se toman muchas muestras de tamaño n de una población muy numerosa de personas y se calcula la calificación media del CI para cada muestra.
 - Si el tamaño de la muestra es $n = 64$, determine la media y la desviación estándar de la distribución de las medias muestrales.
 - Si el tamaño de la muestra es $n = 100$, determine la media y la desviación estándar de la distribución de las medias muestrales.
 - ¿Por qué la desviación estándar del inciso a es diferente de la desviación estándar del inciso b?
- Calificaciones del SAT y el teorema del límite central.** Con base en información del College Board, suponga que las calificaciones del SAT se distribuyen normalmente con una media de 1518 y una desviación estándar de 325. Suponga que se toman muchas muestras de tamaño n de una población grande de estudiantes y se calcula la media de la calificación del SAT para cada muestra.
 - Si el tamaño de la muestra es $n = 100$, determine la media y la desviación estándar de la distribución de las medias muestrales.
 - Si el tamaño de la muestra es $n = 2500$, determine la media y la desviación estándar de la distribución de las medias muestrales.
 - ¿Por qué la desviación estándar en el inciso a es diferente de la desviación estándar del inciso b?
- Dados con doce lados y el teorema del límite central.** Lanzar un dado justo de *doce lados* produce un conjunto de números uniformemente distribuidos entre 1 y 12, con una media de 6.5 y una desviación estándar de 3.452. Suponga que n dados de doce lados se tiran muchas veces y cada vez se registra la media de los n resultados.
 - Determine la media y la desviación estándar de la distribución resultante de las medias de la muestra para $n = 81$.
 - Determine la media y la desviación estándar de la distribución resultante de las medias de la muestra para $n = 100$.
 - ¿Por qué la desviación estándar del inciso a es diferente de la desviación estándar en el inciso b?

8. Dados con diez lados y el teorema del límite central.

Lanzar un dado justo de *diez lados* produce un conjunto de números uniformemente distribuidos entre 1 y 10, con una media de 5.5 y una desviación estándar de 2.872. Suponga que n dados de diez lados se tiran muchas veces y cada vez se registra la media de los n resultados.

- Determine la media y la desviación estándar de la distribución resultante de las medias de la muestra para $n = 49$.
- Determine la media y la desviación estándar de la distribución resultante de las medias de la muestra para $n = 400$.
- ¿Por qué la desviación estándar del inciso a es diferente de la desviación estándar en el inciso b?

Antigüedad de aeronaves. En los ejercicios 9 al 12 suponga que las edades de aeronaves comerciales se distribuyen normalmente con una media de 13.0 años y una desviación estándar de 7.9 años (datos de Servicios de Aviación).

- ¿Qué porcentaje de aeronaves individuales tiene una antigüedad mayor a 15 años? Suponga que seleccionó una muestra aleatoria de 49 aeronaves y calculó la antigüedad media de la muestra. ¿Qué porcentaje de medias muestrales tiene edades mayores que 15 años?
- ¿Qué porcentaje de aeronaves individuales tiene una antigüedad menor a 10 años? Suponga que seleccionó una muestra aleatoria de 84 aeronaves y calculó la antigüedad media de la muestra. ¿Qué porcentaje de medias muestrales tiene edades mayores que 10 años?
- ¿Qué porcentaje de aeronaves individuales tiene una antigüedad entre 10 y 16 años? Suponga que seleccionó una muestra aleatoria de 81 aeronaves y calculó la antigüedad media de la muestra. ¿Qué porcentaje de medias muestrales tiene edades entre 10 y 16 años?
- ¿Qué porcentaje de aeronaves individuales tiene una antigüedad entre 12.5 y 13.5 años? Suponga que seleccionó una muestra aleatoria de 400 aeronaves y calculó la antigüedad media de la muestra. ¿Qué porcentaje de medias muestrales tiene edades entre 12.5 y 13.5 años?
- Cantidades de cola.** Suponga que latas de cola se llenan de modo que la cantidad real se distribuye de manera normal con una media de 12.00 onzas y una desviación estándar de 0.11 onzas.
 - ¿Cuál es la posibilidad de que una muestra de 35 latas tendrá una cantidad media de al menos 12.05 onzas?
 - Dado el resultado en el inciso a, ¿es razonable creer que las latas en realidad se llenan con una media de 12.00 onzas? Si la media no es 12.00 onzas, ¿se está engañando a los consumidores?
- Diseño de luces estroboscópicas.** La luz estroboscópica de una aeronave está diseñada para que los tiempos entre

intermitencia se distribuyan de manera normal con una media de 3.00 segundos y una desviación estándar de 0.40 segundos.

- ¿Cuál es la posibilidad de que un tiempo individual sea mayor que 4.00 segundos?
- ¿Cuál es la posibilidad de que la media de 60 tiempos elegidos aleatoriamente sea mayor que 4.00 segundos?
- Dado que la luz estroboscópica tiene el objetivo de ayudar a otros pilotos a ver una aeronave, ¿cuál resultado es más relevante para la valoración de la seguridad de la luz estroboscópica: el resultado en el inciso a o el resultado en el inciso b? ¿Por qué?

15. Diseño de cascos para motociclistas. Los ingenieros deben considerar la anchura de las cabezas de los hombres cuando diseñan cascos para motociclistas varones. Éstos tienen anchos que se distribuyen de manera normal con una media de 6.0 pulgadas y una desviación estándar de 1.0 pulgada (con base en la encuesta de información antropométrica de Gordon, Churchill, *et al.*).

- Si un hombre es seleccionado aleatoriamente, ¿cuál es la posibilidad de que el ancho de su cabeza sea menor a 6.2 pulgadas?
- La compañía Safeguard Helmet planea una producción inicial de 100 cascos. ¿Cuál es la probabilidad de que 100 hombres elegidos aleatoriamente tengan un ancho medio de cabeza menor que 6.2 pulgadas?
- El gerente de producción ve el resultado del inciso b y argumenta que todos los cascos deben fabricarse para hombres con ancho de cabeza menor que 6.2 pulgadas, ya que les quedarían a casi todos excepto unos cuantos varones. ¿Qué es incorrecto en su razonamiento?

16. Permanencia en agua caliente. Al planear los requerimientos de agua caliente, el gerente del Hotel Luxurion determina que los huéspedes pasan una media de 11.4 minutos del día en la ducha (con base en información de la Compañía de Investigación de Opinión). Suponga que los tiempos de ducha se distribuyen normalmente con una desviación estándar de 2.7 minutos.

- Determine el porcentaje de clientes que se duchan por más de 12 minutos.
- El hotel ha instalado un sistema que puede proporcionar suficiente agua caliente siempre que la media del tiempo de ducha para 84 clientes sea menor que 12 minutos. Si actualmente el hotel tiene 84 clientes, ¿cuán probable es que el hotel no disponga de suficiente agua caliente? ¿El sistema actual parece ser efectivo?

17. Rediseño de asientos expulsores. Cuando se permitió a las mujeres convertirse en pilotos de aviones de guerra, los ingenieros necesitaron rediseñar los asientos de expulsión, ya que éstos habían sido diseñados sólo para hombres. Los asientos de expulsión ACES-II fueron diseñados

para hombres que pesaban entre 140 y 211 libras. El peso de la población de mujeres tiene distribución normal con una media de 143 libras y una desviación estándar de 29 libras (con base en información de la Encuesta Nacional de Salud).

- ¿Qué porcentaje de mujeres tienen pesos entre 140 y 211 libras?
- Si se seleccionan al azar 36 mujeres, ¿cuán probable es que su peso medio esté entre 140 y 211 libras?
- Para el rediseño de los asientos expulsores en aviones de combate, ¿qué es más relevante: el resultado en el inciso a o el resultado en el inciso b? ¿Por qué?

18. Rotulación de paquetes de M&M. Los dulces M&M sencillos tienen pesos que están distribuidos normalmente con un peso medio de 0.8565 gramos y una desviación estándar de 0.0518 gramos (con base en medidas de uno de los autores). Una muestra aleatoria de 100 M&M se obtiene de un paquete con 465 dulces; la etiqueta del paquete indica un peso neto de 396.9 gramos. (Si cada paquete tiene 465 dulces, el peso medio de los dulces debe exceder a $396.9/465 = 0.8535$ para que el contenido neto pese al menos 396.9 gramos).

- Si un dulce M&M sencillo se elige aleatoriamente, ¿qué posibilidad hay que pese más de 0.8535?
- Si 465 dulces M&M se eligen aleatoriamente, ¿qué posibilidad hay que su peso medio sea al menos 0.8535 gramos?
- Dados estos resultados, ¿parece que la compañía Mars está proporcionando a los consumidores de M&M con la cantidad que asegura en la etiqueta?

19. Máquinas expendedoras. En la actualidad los cuartos de dólar tienen pesos que se distribuyen normalmente con una media de 5.670 gramos y una desviación estándar de 0.062 gramos. Una máquina expendedora está configurada para aceptar sólo aquellos cuartos de dólar con pesos entre 5.550 y 5.790 gramos.

- Si se insertan 280 cuartos de dólar diferentes en la máquina expendedora, ¿cuál es el número esperado de cuartos de dólar rechazados?
- Si se insertan 280 cuartos de dólar diferentes en la máquina expendedora, ¿cuál es la probabilidad de que la media caiga entre los límites de 5.550 y 5.790 gramos?
- Si usted es el propietario de la máquina expendedora, ¿cuáles resultados le interesarían más, el resultado del inciso a o el resultado del inciso b? ¿Por qué?

20. Normas de seguridad en aeronaves. Bajo de las antiguas reglas de la Administración Federal de Aviación, las aerolíneas habían estimado los pesos de los pasajeros como

185 libras. (Esa cantidad es para un adulto que viaja en invierno, e incluye 20 libras de equipaje de mano). Las reglas actuales requieren una estimación de 195 libras. Los hombres tienen pesos que se distribuyen normalmente con una media de 172 libras y una desviación estándar de 29 libras.

- Si un hombre adulto es seleccionado al azar y se supone que lleva 20 libras de equipaje de mano, ¿cuál es la probabilidad de que su peso total sea mayor a 195 libras?
- Si un avión Boeing 767-300 está lleno con 213 pasajeros adultos hombres y cada uno lleva 20 libras de equipaje de mano, ¿qué posibilidad hay de que el peso medio de los pasajeros (incluyendo su equipaje de mano) sea mayor que 195 libras? ¿Un piloto tiene que estar preocupado por el exceso de este peso límite?

21. Variable genérica. Suponga una variable, tal como estatura o presión arterial, que está distribuida normalmente en una población grande con una media de μ y una desviación estándar de σ .

- Suponga que tomamos muchas muestras de tamaño $n = 100$ y calculamos la media de la variable para cada muestra. ¿Cómo se compara la desviación estándar de la distribución de las medias muestrales con σ ? Explique por qué.
- Suponga que tomamos muchas muestras de tamaño $n = 1000$ y calculamos la media de la variable para cada muestra. ¿Cómo se compara la desviación estándar de esta distribución de medias muestrales con la desviación estándar de la distribución de medias muestrales del inciso a? ¿Cómo se compara la desviación estándar de esta distribución de medias muestrales con σ ? Explique sus respuestas.
- ¿Cómo cambia la desviación estándar de la distribución de las medias muestrales conforme tomamos muestras de tamaño cada vez mayores?

22. Presión arterial en mujeres. La presión arterial sistólica para mujeres entre 18 y 24 años se distribuye normalmente con una media de 114.8 (milímetros de mercurio) y una desviación estándar de 13.1.

- ¿Cuál es la posibilidad de que la presión arterial de una mujer individual esté por arriba de 125?
- Suponga que se selecciona una muestra aleatoria de $n = 300$ mujeres y se calcula la media de la presión arterial de la muestra. ¿Cuál es la probabilidad de que la media de la presión arterial para la muestra será superior a 125?
- Suponga que se selecciona una muestra aleatoria de $n = 300$ mujeres y calcula la media de la presión arterial de la muestra. ¿Cuál es la probabilidad de que la media de la presión arterial para la muestra será inferior a 114?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 5 en www.aw.com/bbt.

- 23. Teorema del límite central en internet.** Haciendo una búsqueda en internet sobre “teorema del límite central” descubrirá muchos sitios dedicados a este tema. Encuentre un sitio que tenga demostraciones animadas del teorema del límite central. Describa con sus palabras lo que observe y cómo se ilustra el teorema del límite central.
- 24. El quincunx en internet.** Haga una búsqueda en internet sobre el “teorema del límite central” o de “quincunx” y en-

cuentre un sitio que tenga una demostración animada del quincunx (o tablero de Galton). Describa el quincunx y explique cómo ilustra el teorema del límite central.

- 25. Tiro de dados.** Demuestre el teorema del límite central usando dados, como se estudió en esta sección. Proporcione a cada persona en su clase tantos dados como sea posible. Inicie lanzando un dado y construyendo un histograma de los resultados. Luego cada persona tira dos dados y construye un histograma de la media de cada tiro. Aumente el número de dados en cada tiro, tantos como dados y tiempo sea posible. Comente sobre las apariencias del histograma en cada etapa.

Ejercicios de repaso del capítulo

1. Para cada una de las situaciones siguientes, indique si la distribución de valores es probable que tenga una distribución normal. Proporcione una explicación breve que justifique su elección.
 - a. Números que resultan de girar una ruleta. (Hay 38 ranuras igualmente probables con números 0, 00, 1, 2, 3, ..., 36).
 - b. Pesos de perros adultos Golden Retriever.
 - c. Medidas de tiempos de reacción al frenado de conductores de 18 años de edad.
2. Las calificaciones en el examen ACT se distribuyen normalmente con una media de 21.1 y una desviación estándar de 4.8.
 - a. Con la regla 68-95-99.7, determine el porcentaje de calificaciones del ACT dentro de 9.6 de la media, 21.1.
 - b. Con la regla 68-95-99.7, determine el porcentaje de calificaciones ACT dentro de 14.4 de la media, 21.1.
 - c. Una calificación de 35.6, ¿es rara? ¿Por qué sí o por qué no?
3. Suponga que la temperatura corporal de adultos sanos se distribuye normalmente con una media de 98.20°F y una desviación estándar de 0.62°F (con base en información de investigadores de la Universidad de Maryland).
 - a. Si usted tiene una temperatura corporal de 99.00°F , ¿cuál es su puntuación percentil?
 - b. Convierta 99.00°F a una puntuación estándar (o puntuación z).
 - c. Una temperatura corporal de 99.0°F , ¿es rara? ¿Por qué sí o por qué no?
 - d. Cincuenta adultos se eligen aleatoriamente. ¿Cuál es la posibilidad de que la media de sus temperaturas corporales sea 97.98°F o menor?
 - e. La temperatura corporal de una persona se encuentra que es 101.00°F . ¿Este resultado es “inusual”? ¿Por qué sí o por qué no? ¿Qué debe concluir?
 - f. ¿Qué temperatura corporal es el percentil 95o?
 - g. ¿Qué temperatura corporal es el quinto percentil?
 - h. El hospital Bellevue en la ciudad de Nueva York utiliza 100.6°F como la temperatura mínima considerada para indicar fiebre, ¿Qué porcentaje de adultos normales y sanos se considerarían con fiebre? ¿Este porcentaje sugiere que un corte de 100.6°F es apropiado?
 - i. Si, en lugar de suponer que la temperatura media del cuerpo es 98.20°F , suponemos que la media es 98.60°F (como mucha gente cree), ¿cuál es la posibilidad de que al seleccionar aleatoriamente a 106 personas y tomarles la temperatura se obtenga una media de 98.20°F o inferior? (Continúe suponiendo que la desviación estándar es 0.62°F). Los investigadores de la Universidad de Maryland obtuvieron tal resultado. ¿Qué debe concluir?

Cuestionario del capítulo

1. ¿Cuáles de los enunciados siguientes son correctos?
 - a. Una distribución normal es cualquier distribución que no sea inusual.
 - b. La gráfica de una distribución normal tiene forma de campana.
 - c. Si una población tiene una distribución normal, la media y la median no son iguales.
 - d. En una distribución normal es posible que la desviación estándar sea 0.
 - e. La gráfica de una distribución normal es simétrica.
2. Una prueba de percepción está diseñada de modo que las calificaciones se distribuyen normalmente con una media de 50 y una desviación estándar de 10. Usando la regla 68-95-99.7, determine el porcentaje de calificaciones dentro de 20 puntos de la media de 50.
3. Una población se distribuye normalmente con una media de 50 y se seleccionan muestras aleatorias de tamaño 100. ¿Cuál es la media de las medias muestrales?
4. Una población se distribuye normalmente con una desviación estándar de 5 y se seleccionan muestras aleatorias de tamaño 100. ¿Cuál es la desviación estándar de las medias muestrales?
5. Una población tiene una distribución normal con una media de 50 y una desviación estándar de 10. ¿Cuál es la puntuación z que corresponde a 70?
6. Una población tiene una distribución normal con una media de 50 y una desviación estándar de 10. ¿Cuál es la puntuación z que corresponde a 40?
7. Una población de calificaciones de exámenes tiene una distribución normal con una media de 50 y una desviación estándar de 10. ¿Qué porcentaje de calificaciones son mayores que 50?

8. Una población de calificaciones de un examen tiene una distribución normal con una media de 50 y una desviación estándar de 10. Si 97.72% de las calificaciones son menores que 70, ¿qué porcentaje de calificaciones son mayores que 70?
9. Una población de calificaciones de un examen tiene una distribución normal con una media de 50 y una desviación estándar de 10. Si 97.72% de las calificaciones son menores que 70, ¿qué porcentaje de calificaciones son mayores que 30?
10. ¿Cuál de las siguientes es probable que tenga una distribución que sea cercana a una distribución normal?
 - a. Los resultados que ocurren cuando un solo dado se tira muchas veces.
 - b. Los resultados que ocurren cuando dos dados se tiran muchas veces y la media se calcula cada vez.
 - c. Los resultados que ocurren cuando cinco dados se tiran muchas veces y la media se calcula cada vez.

Uso de la tecnología

La tabla 5.1 incluye percentiles que corresponden a valores limitados de puntuaciones estándar z , pero los paquetes estadísticos de cómputo por lo regular permiten el uso de cualquier puntuación estándar z . La tabla 5.1 lista percentiles basados en valores acumulados a la izquierda, por lo común los paquetes de cómputo tienen el mismo formato, pero utilizan áreas en lugar de percentiles. Por ejemplo, la tabla 5.1 muestra el percentil acumulado a la izquierda de 97.72 para $z = 2.0$, pero los paquetes darán un área acumulada a la izquierda (o probabilidad) de 0.9772.

Entre los programas analizados a continuación, STATDISK es por mucho el más sencillo de usar. STATDISK o cualquier otro programa adecuado reemplaza a la tabla 5.1 y le proporciona un rango mucho más amplio de valores que pueden usarse.

SPSS

- Para determinar el área acumulada (o probabilidad) a la izquierda de un valor particular, siga este procedimiento:
 1. Ingrese un valor x de en la ventana del editor de datos de SPSS y luego presione **ENTER**.
 2. Dé clic en el elemento de menú **Transform** en la ventana del editor de datos de SPSS.
 3. Dé clic en el elemento **Compute**.
 4. En la ventana “Compute Variable” haga estas dos entradas:
 - a. Ingrese **prob** en el cuadro “Target Variable”.
 - b. Ingrese **CDF.NORMAL(x, media, desviación estándar)** en la caja “Numeric Expression”, donde se utilizan valores específicos para cada una de las tres entradas. (Si utiliza puntuaciones z , ingrese una media de 0 y una desviación estándar de 1).
 5. Dé clic en **OK**.
 6. La probabilidad (o área) correspondiente a la región a la izquierda del valor x se mostrará. Dé clic en ese valor para ver más lugares decimales.
- Para determinar el valor de x que corresponda a un área acumulada a la izquierda, siga este procedimiento:
 1. En la ventana del editor de datos de SPSS ingrese el valor para el área acumulada (o probabilidad) que corresponde a la región a la izquierda del valor deseado x , y luego presione **ENTER**.

2. Dé clic en el elemento del menú **Transform** en la ventana del editor de datos de SPSS.
3. Dé clic en el elemento **Compute**.
4. En la ventana “Compute Variable”, haga estas dos entradas:
 - a. Ingrese x en el cuadro “Target Variable”.
 - b. Ingrese **IDF.NORMAL(p, media, desviación estándar)** en el cuadro “Numeric Expression”, donde se utilizan valores específicos para cada una de las tres entradas. El valor de p debe ser el área acumulada a la izquierda del valor deseado. (Si utiliza puntuaciones estándar z , ingrese una media de 0 y una desviación estándar de 1).
5. Dé clic en **OK**.
6. El valor de x se mostrará. Dé clic en ese valor para ver más lugares decimales.

Excel

- Para determinar el área acumulada a la izquierda de un valor (tal como 2), dé clic en **fx** y luego, en el cuadro de diálogo, seleccione **Estadísticas, DISTR.NORM (Statistical, NORMDIST)**. En el cuadro de diálogo ingrese el valor (tal como 2) para x , ingrese la media y la desviación estándar, e ingrese 1 en el espacio “acum” (“cumulative”). (Si utiliza puntuaciones estándar z , ingrese una media de 0 y una desviación estándar de 1, e ingrese 1 en el espacio “acum”).
- Para determinar un valor que corresponda a un área conocida (tal como 0.95) seleccione **fx, Estadísticas, DISTR.NORM.INV (Statistical, NORMINV)**, y proceda a hacer las entradas en el cuadro de diálogo. (Si utiliza puntuaciones estándar, ingrese una media de 0 y una desviación estándar de 1). Cuando ingrese el valor de la probabilidad, ingrese el área total a la izquierda del valor dado.

STATDISK

Seleccione **Analysis, Probability Distributions, Normal Distribution**. Ingrese la puntuación estándar z para encontrar el área correspondiente o bien introduzca el área acumulada a la izquierda para determinar la puntuación z . Después de introducir un valor, dé clic en **Evalue**.

HABLEMOS DE EDUCACIÓN



¿Qué podemos aprender de las tendencias del SAT?

El Examen de Aptitud Escolar (SAT, por sus siglas en inglés) ha sido tomado por estudiantes que terminan los estudios preuniversitarios desde 1941, cuando 11 000 estudiantes presentaron por primera vez el examen. En la actualidad, cada año, lo presentan aproximadamente 1.5 millones de estudiantes preuniversitarios, y lo presentan en alguna ocasión casi la mitad de todos los alumnos que terminan la preparatoria. Hasta 2006, había dos partes en el SAT: la prueba verbal y la prueba de matemáticas. En 2006 el examen verbal se cambió para convertirlo en examen de “lectura de comprensión” y se agregó un tercer examen sobre redacción. También se hicieron cambios en el examen de matemáticas.

Originalmente, las calificaciones en cada parte del SAT fueron graduadas para tener una media de 500 y una desviación estándar de 100. La calificación máxima y mínima en cada parte son 800 y 200 respectivamente, que son 3 desviaciones arriba y abajo de la media. (Calificaciones de más de tres desviaciones estándar por arriba o por abajo de la media sólo obtienen el máximo 800 o el mínimo 200).

Sin embargo, la media no permanece en 500 año con año. Cada examen del año comparte algunas preguntas comunes con los exámenes de años anteriores, y el College Board (Consejo Escolar, que diseña el examen) utiliza estas preguntas comunes para comparar cada año los resultados con los de años anteriores. Por tanto, el College Board asegura que una calificación de 500 represente el mismo nivel de logro en el examen sin importar en qué año se presentó. Por ejemplo, si la calificación media en un año particular es superior a 500, implica que los estudiantes se desempeñaron mejor en promedio que los estudiantes en el que el promedio fue originalmente tomado como 500.

Las tendencias en las calificaciones del SAT han sido ampliamente utilizadas para evaluar el estado general de la educación en Estados Unidos. La figura 5.27 muestra las calificaciones promedio en las partes verbal/lectura de comprensión y matemáticas del SAT entre 1972 y 2006; el rótulo “verbal/lectura de comprensión” se utiliza porque se refiere al examen verbal antes de 2006 y al examen de lectura de comprensión iniciado en 2006. Observe que las calificaciones para ambas partes por lo general disminuyen hasta principios de los noventa, pero se han recuperado en años recientes. Los resultados para los exámenes de redacción no se muestran, porque no hay suficiente información para observar tendencias sobre el tiempo.

Las tendencias muestran que, si las calificaciones de año con año en realidad son comparables, entonces los estudiantes que presentan el SAT en años recientes tienen habilidades verbales más pobres que los estudiantes de hace unas cuantas décadas, y las calificaciones combinadas también son un poco menores. Pero antes de que aceptemos esa conclusión, debemos responder dos preguntas importantes:

1. ¿Las tendencias entre la *muestra* de estudiantes de preparatoria quienes presentan el SAT son representativas de las tendencias para la población de todos los estudiantes de preparatoria?
2. Además del pequeño número de preguntas comunes que relacionan un año con el siguiente, el examen cambia cada año. ¿Las calificaciones de un año se pueden comparar legítimamente con las de otros años?

Por desgracia, ninguna pregunta puede responderse con un rotundo sí. Por ejemplo, muchas personas argumentan que gran parte de la disminución en las calificaciones del SAT es el resultado de cambios en la muestra de estudiantes que presentan el examen. En la década de

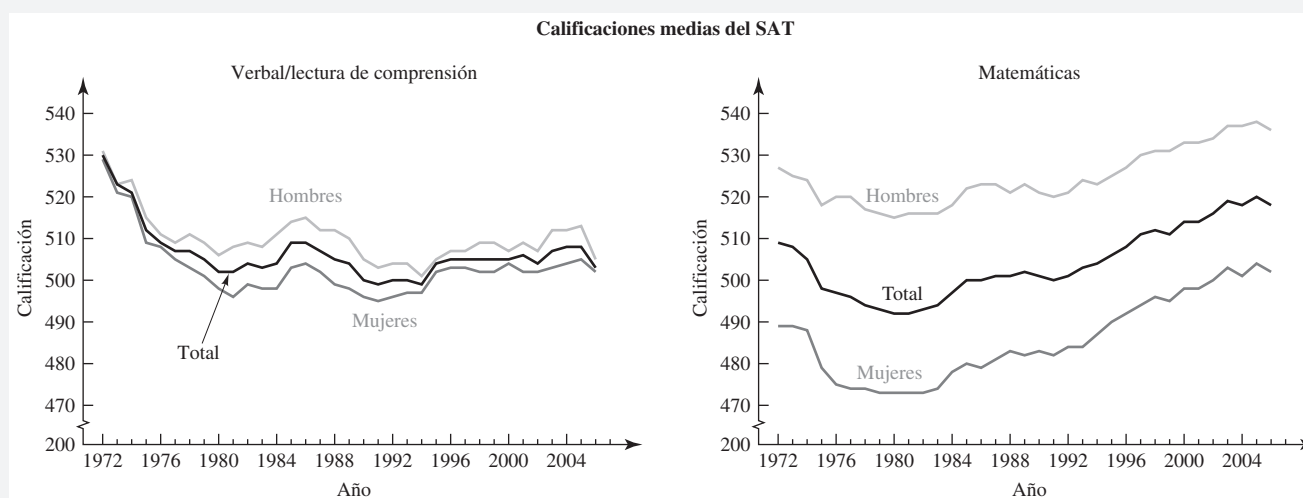


Figura 5.27 Calificaciones en el área verbal/lectura de comprensión y matemáticas del SAT, 1972-2006. Las calificaciones para los años anteriores a 1996 se muestran en sus valores “re-centrados”, en lugar de sus valores originales. *Fuente:* The College Board (Consejo de Educación, Estados Unidos).

los setenta sólo una tercera parte de los estudiantes que terminaron la preparatoria presentaron el SAT, mientras que casi 50% de los estudiantes que terminaron la preparatoria presentaron el SAT en 2006. Si estas muestras representan el nivel superior de los estudiantes de preparatoria, la disminución en los promedios de calificaciones podría reflejar simplemente el hecho de que un rango más amplio de capacidades está representado en las últimas muestras. El respaldo para esta idea proviene de la información estado por estado del SAT. En general, los estados con la media más alta de calificaciones del SAT también son aquéllos con las proporciones más bajas de sus poblaciones que presentan el examen. La explicación que se da para esta tendencia —una explicación que también está apoyada por un análisis estadístico más detallado— es que cuando una proporción más pequeña de estudiantes de un estado presenta el examen, esos estudiantes tienden a estar entre los mejores estudiantes de preparatoria del estado.

La segunda pregunta es aún más difícil. Incluso si algunas preguntas siguen siendo las mismas de un año al siguiente, diferente material podría enfatizarse en las preparatorias, así que cambia la posibilidad de que los estudiantes respondan correctamente las mismas preguntas de un año al siguiente.

Más complicaciones surgen cuando los exámenes sufren mayores cambios. Por ejemplo, en 1994 el uso de calculadoras fue permitido para la primera parte en la sección de matemáticas, y tiempo adicional se permitió para todo el examen (el examen verbal también sufrió cambios significativos ese año). ¿Es una coincidencia que las calificaciones de matemáticas empezaran una tendencia a la alza ese año, o podría ser que el examen se hizo “más sencillo” como resultado de los cambios? De manera similar, en 2006 las calificaciones combinadas de lectura de comprensión y matemáticas sufrieron las disminuciones más grandes en más de 30 años. Los críticos atribuyen la disminución al hecho de que el SAT se hizo más extenso y difícil como resultado de la adición del examen de redacción, mientras que el College Board atribuyó la disminución a una baja en el número de estudiantes que volvían a presentar el examen. (Los estudiantes que vuelven a presentar el examen tienden a elevar la calificación promedio, ya que obtienen un aumento significativo en sus calificaciones cuando presentan el examen una segunda o tercera vez). Es claro que las modificaciones en la estructura del examen hacen difícil determinar si las alteraciones en el promedio de calificaciones reflejan cambios reales en la educación.

Quizás aún más significativo, la calibración del examen del SAT sufrió un cambio importante en 1996. En ese tiempo las calificaciones habían disminuido de manera significativa respecto a la media planeada de 500: la calificación media verbal había caído a alrededor de 420, mientras que la calificación media en matemáticas había caído a alrededor de 470. Puesto que el rango de calificaciones posibles va sólo de 200 a 800, estas disminuciones significaban que las calificaciones mínima y máxima ya no representarían el mismo número de

desviaciones estándar de la media, lo cual creó dificultades con el análisis estadístico. El College Board decidió “re-centrar” todas las calificaciones a una media de 500 en 1996. Esto afectó de forma diferente a percentiles diferentes, pero efectivamente agregó 80 puntos a la media de las calificaciones de la parte verbal y 30 puntos a la media de las calificaciones de matemáticas. Las calificaciones mostradas en la figura 5.27 son las calificaciones “re-centradas”, por lo que para todos los años anteriores a 1996, son alrededor de 80 puntos más que las calificaciones reales para el examen verbal y alrededor de 30 puntos más que las calificaciones reales para el examen de matemáticas. Por ejemplo, la figura 5.27 muestra una calificación verbal media de casi exactamente 500 en 1994, pero si usted regresa a los reportes de ese tiempo, verá que la media real fue alrededor de 420 en 1994.

El College Board argumenta que su análisis estadístico ha sido realizado con gran cuidado, y que a pesar de volver a centrar y los cambios en el examen, las tendencias en el SAT reflejan cambios reales en la educación. Los críticos en general coinciden en que el análisis estadístico está bien hecho, pero argumentan que otros factores, tal como el cambio en la población que presenta el examen y cambios en nuestro sistema educativo, hacen que no sean válidas las comparaciones. Sin embargo, para los estudiantes el hecho clave es que muchas universidades siguen utilizando las calificaciones del SAT para tomar sus decisiones sobre la admisión de los solicitantes. Mientras permanezca el caso, el SAT continuará como uno de los exámenes más importantes en Estados Unidos, y el debate sobre sus méritos seguramente continuará.

PREGUNTAS PARA DISCUSIÓN

1. ¿Usted considera que las comparaciones de calificaciones del SAT en dos años tienen sentido? ¿En diez años? ¿Coincide en que las tendencias a largo plazo indican que los estudiantes de hoy tienen peores habilidades verbales pero mejores habilidades matemáticas que los de hace unas cuantas décadas? Defienda sus opiniones.
2. La tendencia a la baja en habilidades verbales fue una razón principal para que el College Board agregara el examen de redacción en 2006. La esperanza era que al agregar este nuevo examen causaría que los estudiantes se centrasen más en la escritura en la preparatoria y por tanto mejorasen tanto en redacción como en habilidades verbales. ¿Considera que esto es racional? ¿Considera que tendrá éxito?
3. Observe que las calificaciones para hombres han sido consistentemente más altas que las de mujeres. ¿Por qué cree que sea el caso? ¿Cree que cambios en nuestro sistema educativo podrían eliminar esta brecha? Defienda sus opiniones.
4. Analice experiencias personales con el SAT entre sus compañeros. Con base en estas experiencias personales, ¿qué está midiendo el SAT? ¿Considera que el examen es una manera razonable de predecir el desempeño de estudiantes en la universidad? ¿Por qué sí o por qué no?

LECTURAS RECOMENDADAS

Arenson, Karen W., “Scores on Reading and Math Portions of SAT Show Significant Decline”, *New York Times*, 30 de agosto de 2006.

College Board, “The Effects of SAT Scale Recentering on Percentiles”, Resumen de investigación RS-05, 1999, disponible en http://www.collegeboard.com/repository/rs05_3962.pdf.

Crouse, James y Trusheim, Dale, *The Case Against the SAT*, University of Chicago Press, 1988.

HABLEMOS DE PSICOLOGÍA

¿Somos más inteligentes que nuestros padres?

La mayoría de los niños tienden a pensar que son más inteligentes que sus padres, pero ¿es posible que esto sea cierto? Si creemos los resultados de pruebas de CI, no sólo somos más inteligentes que nuestros padres, en promedio, sino que nuestros padres son más inteligentes que nuestros abuelos. De hecho, casi todos nosotros estaríamos catalogados como genios si regresáramos cien años al pasado. Por supuesto, antes de que cualquiera de nosotros comience a promover nuestras habilidades casi de Einstein, sería bueno investigar qué está atrás de esta sorprendente afirmación.

La idea de un CI, que se establece por *cociente de inteligencia*, fue inventada por el psicólogo francés Alfred Binet (1857-1911). Binet creó un examen para identificar a los niños con necesidad de apoyo escolar especial. Él aplicó su examen a muchos niños, y luego calculó cada CI del niño dividiendo la “edad mental” del niño entre su edad psicológica (y multiplicando por 100). Por ejemplo, un niño de 5 años que obtuvo una calificación como un niño promedio de 6 años decía que tenía una edad mental de 6, y por tanto un CI de $(6 \div 5) \times 100$ o 120. Observe que, por esta definición, los exámenes de CI sólo tienen sentido para niños. Sin embargo, los investigadores posteriores, en especial los psicólogos del ejército de Estados Unidos, extendieron la idea del CI de modo que pudiese aplicarse también a adultos.

Actualmente, el CI se define mediante una distribución normal con una media de 100 y una desviación estándar de 16. De acuerdo con la regla 68-95-99.7, 68% de la gente que toma un examen de CI obtiene calificaciones entre 84 y 116, 95% obtiene calificaciones entre 68 y 132 y 99.7 obtiene calificaciones entre 52 y 148. Por tradición los psicólogos clasificaron a la gente con un CI debajo de 70 (alrededor de 2 desviaciones estándar por debajo de la media de 100 como “intelectualmente deficientes” y gente con una calificación por arriba de 130 (alrededor de dos desviaciones estándar arriba de la media) como “intelectualmente superior”.

Quizás esté consciente de la controversia que rodea a los exámenes de CI, la cual se reduce a dos temas clave:

- ¿El CI mide la inteligencia o alguna otra cosa?
- Si mide la inteligencia, ¿es algo que es innato y determinado por herencia o algo que puede moldearse por el ambiente y la educación?

Una discusión completa de este tema es demasiado complicada para tratarla aquí, pero un descubrimiento reciente descubrió una tendencia en las calificaciones del CI que arrojaron luz en estos temas de una manera sorprendente. Como veremos en breve, la tendencia es muy pronunciada, pero estuvo oculta por la manera en que se calificaba el CI. Existen varias versiones de exámenes del CI, y la mayoría de ellas se cambian con regularidad y se actualizan. Pero en todos los casos las calificaciones se ajustan a una distribución normal con una media de 100 y una desviación estándar de 16. En otras palabras, la calificación de un examen de CI en esencia se realiza de la misma manera que un instructor podría clasificar un examen “sobre una curva”. A consecuencia de este ajuste, la media en los exámenes de CI *siempre* es 100, lo cual hace imposible medir calificaciones del CI que suban o bajen con el tiempo.

Sin embargo, unos cuantos exámenes del CI no han cambiado y actualizado al paso del tiempo, incluyendo algunos para los militares. En otros casos, exámenes que han sido actualizados algunas veces aún repiten preguntas antiguas. A principios de los ochenta del siglo pasado, un profesor de ciencias políticas, llamado doctor James Flynn, empezó a buscar las



A propósito...

Binet mismo supuso que la inteligencia podría ser moldeada y advirtió de tomar sus exámenes como una medida de cualquier habilidad innata o heredada. Sin embargo, psicólogos muy posteriores concluyeron que los exámenes de CI podrían medir inteligencia innata, lo que condujo a su utilización para separar niños en edad escolar, reclutas militares y muchos otros grupos de personas de acuerdo con una supuesta habilidad intelectual.

calificaciones sin procesar y sin ajustar de exámenes y preguntas que no han cambiado. Los resultados fueron asombrosos.

El doctor Flynn encontró que las calificaciones sin procesar se han mantenido constantemente a la alza, aunque la cantidad precisa de aumento varía un poco con el tipo de examen de CI. Las tasas más altas de aumento se encontraron en exámenes que pretendían medir habilidades de razonamiento abstracto (tales como pruebas “Raven”). Para estas pruebas, el doctor Flynn encontró que las calificaciones no ajustadas del CI de personas en países industrializados se han elevado a una tasa de alrededor de 6 puntos por década. En otras palabras, una persona que obtuvo 100 en una prueba en el año 2000 habría obtenido alrededor de 106 en un examen de 1990, 112 en un examen de 1980, y así sucesivamente. A lo largo de un siglo esto implicaría un aumento de 60 puntos, sugiriendo que alguien con clasificación del CI de “intelectualmente deficiente” de 70, ahora habría sido catalogado con un CI de 130, “intelectualmente superior”, hace un siglo.

Esta tendencia a largo plazo de elevación de las puntuaciones en exámenes del CI ahora se conoce como *efecto Flynn*. Está presente para todo tipo de exámenes de CI, aunque no siempre con el mismo grado que los exámenes de razonamiento abstracto. Por ejemplo, la figura 5.28 muestra cómo cambiaron los resultados en uno de los exámenes de CI más ampliamente usados (el examen Stanford-Binet) entre 1932 y 1997. Observe que en términos de calificaciones no ajustadas, la media se elevó 20 puntos en ese periodo. En otras palabras, personas que obtuvieron un CI de 100 en un examen de 1997 hubiesen sido clasificadas con un CI de 120 en un examen de 1932. Como lo muestran las cifras, alrededor de un cuarto de los que presentaron examen en 1997 serían clasificados como “intelectualmente superiores” en el examen de 1932. Existe alguna evidencia de que el aumento en las calificaciones puede haber empezado a hacerse más lento o detenido en años recientes, aunque la información aún está sujeta a debate.

Muchos otros científicos han investigado el efecto Flynn, y parece que hay un poco de duda de que la tendencia a largo plazo sea real. La implicación es clara: lo que sea que midan los exámenes de CI, la gente actualmente en realidad *tiene* más de él que la gente de hace unas cuantas décadas. Si los exámenes del CI miden la inteligencia, entonces significa que en realidad somos más inteligentes que nuestros padres (en promedio), quienes a su vez son más inteligentes que nuestros abuelos (en promedio).

Por supuesto, si los exámenes del CI no miden “inteligencia” sino sólo miden algún tipo de habilidad, entonces el aumento en las calificaciones puede indicar que los niños de hoy tienen más práctica en esa habilidad que los niños de ayer. El hecho de que el mayor aumento sea visto en pensamiento abstracto conduce a apoyar esta idea. Estos exámenes incluyen problemas tales como resolver rompecabezas y buscar patrones entre conjuntos de formas y estos tipos de problemas ahora son mucho más comunes en los juegos que lo que fueron en el pasado.

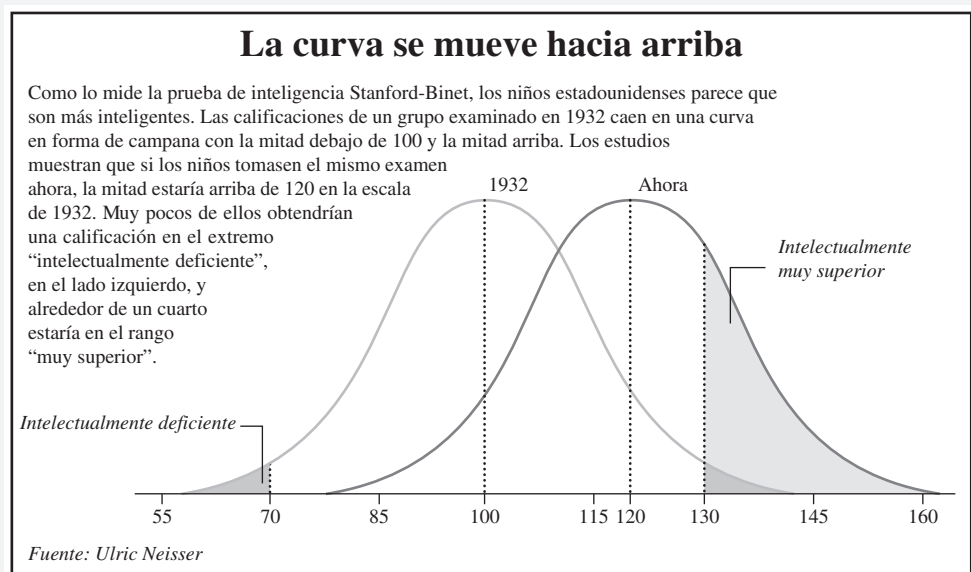


Figura 5.28 Fuente: New York Times.

Aunque el efecto Flynn no responde si los exámenes del CI miden inteligencia, pueden decirnos una cosa importante: si los CI en realidad se han elevado como lo sugiere el efecto Flynn, entonces el CI *no* es un rasgo heredado, ya que los rasgos heredados no pueden cambiar mucho en solo unas cuantas décadas. Por tanto, si los exámenes de CI están midiendo inteligencia, entonces la inteligencia puede moldearse mediante factores del entorno. Recíprocamente, si la inteligencia es heredada, entonces los exámenes del CI no la miden.

El descubrimiento del doctor Flynn ya ha cambiado la manera como los psicólogos ven los exámenes del CI, y asegura ser un tema de investigación en las próximas décadas. Además, dados los muchos usos para los cuales la sociedad moderna ha puesto los exámenes del CI, el efecto Flynn probablemente también tenga profundas consecuencias sociales y políticas. Así que regresando a nuestra pregunta inicial: ¿Somos más inteligentes que nuestros padres? En realidad no podemos responder, pero ciertamente desearíamos que fuese así, ya que se requerirá de mucha capacidad intelectual para resolver los problemas del futuro.

PREGUNTAS PARA DISCUSIÓN

1. ¿Con cuál explicación del efecto Flynn está a favor: que la gente sea más inteligente o que la gente sólo tiene más práctica en las habilidades que miden los exámenes del CI? Defienda su opinión.
2. El aumento en el desempeño de los exámenes del CI contrasta fuertemente con una disminución sostenida en el desempeño, en las pasadas décadas, en muchos exámenes que miden conocimiento factual, tal como el SAT. Piense en varias formas posibles que expliquen estos resultados contrastantes y forme una opinión como la explicación más probable.
3. Los resultados en exámenes del CI tienden a diferir entre diferentes grupos étnicos. Algunas personas han utilizado este hecho para argumentar que algunos grupos étnicos tienden a ser intelectualmente superiores que otros. ¿Tal argumento aún puede ser sostenido a la luz del efecto Flynn? Defienda su opinión.
4. Analice algunos usos comunes de los exámenes del CI. ¿Usted considera que los exámenes del CI *deben* usarse para estos propósitos? ¿El efecto Flynn altera sus opiniones acerca de los usos de los exámenes del CI? Explique.

LECTURAS RECOMENDADAS

Hall, Trish, "I.Q.: Scores Are Up, and Psychologists Wonder Why", *New York Times*, 24 de febrero de 1998.

Neisser, Ulric (ed.), *The Rising Curve: Long-Term Gains in IQ and Related Measures*, American Psychological Association, 1998.



*Usted puede tomar como
entendido que su suerte
cambia sólo si es bueno.*

—Ogden Nash,
Roulette Us Be Gay

Probabilidad en estadística

LA MAYORÍA DE LOS ESTUDIOS ESTADÍSTICOS BUSCAN aprender algo acerca de una *población* a partir de una *muestra* mucho más pequeña. Por tanto, una cuestión clave en cualquier investigación estadística es la validez de generalizar a partir de una muestra a la población. Para responder esta pregunta debemos entender la verosimilitud, o probabilidad, de que lo aprendido acerca de la muestra también aplica a la población. En este capítulo pondremos atención a unas ideas básicas de probabilidad que se utilizan comúnmente en estadística. Como verá, estas ideas de probabilidad también tienen muchas aplicaciones por sí mismas.

OBJETIVOS DE APRENDIZAJE

- 6.1 El papel de la probabilidad en estadística: significancia estadística**
Entender el concepto de significancia estadística y el papel esencial que desempeña la probabilidad al definirla.
- 6.2 Fundamentos de probabilidad**
Conocer cómo encontrar probabilidades usando los métodos teórico y de frecuencia relativa y entender cómo construir distribuciones básicas de probabilidad.
- 6.3 Probabilidad con grandes números**
Entender la ley de los grandes números, usar esta ley para comprender y calcular valores esperados, y reconocer cómo una mala interpretación de esta ley lleva a la falacia del jugador.
- 6.4 Ideas de riesgo y esperanza de vida**
Calcular e interpretar varias medidas de riesgo aplicadas a viajes, enfermedades y esperanza de vida.
- 6.5 Combinación de probabilidades (sección complementaria)**
Distinguir entre eventos independientes y dependientes y entre eventos traslapados y no traslapados, y ser capaz de calcular probabilidades “y” y probabilidades “o”.



6.1 El papel de la probabilidad en estadística: significancia estadística

Para ver por qué la probabilidad es tan importante en estadística, iniciemos con un ejemplo sencillo del lanzamiento de una moneda. Suponga que usted trata de probar si una moneda es justa, esto es, es igualmente probable que caiga en cara (H) o en cruz (T). Si lanza la moneda 100 veces y obtiene 52 caras y 48 cruces, ¿debe concluir que la moneda no es justa? No. Aunque debe esperar ver *aproximadamente* 50 caras y 50 cruces en cada 100 lanzamientos de una moneda justa, también debe esperar alguna variación de una muestra de 100 lanzamientos a otra. Una pequeña desviación de una perfecta división 50-50 entre caras y cruces no necesariamente significa que la moneda no es justa, ya que esperamos que ocurran pequeñas desviaciones *por el azar*.

Ahora suponga que lanza una moneda 100 veces y los resultados son 20 caras y 80 cruces. Ésta es una desviación significativa de una división 50-50, haciendo parecer mucho menos probable que fue sólo un conjunto al azar de lanzamientos de una moneda justa. En otras palabras, aunque es *posible* que observe un conjunto raro de 100 lanzamientos, es más probable que la moneda no sea justa. Cuando la diferencia entre lo que se observa y lo que se espera parece poco probable para explicarse sólo por el azar, decimos que la diferencia es **estadísticamente significativa**.

Definición

Un conjunto de medidas u observaciones en una investigación estadística es **estadísticamente significativo** si es poco probable que haya ocurrido por el azar.

EJEMPLO 1 ¿Relatos probables?

- a. *Un detective en Detroit encuentra que 25 de las 62 armas usadas en crímenes durante la semana pasada fueron vendidas en la misma armería.* Este hallazgo es estadísticamente significativo. Ya que existen muchas tiendas de armas en el área de Detroit, al tener 25 de 62 armas provenientes de la misma tienda parece poco probable que haya ocurrido por azar.
- b. *En términos de la temperatura global promedio, cinco de los años entre 1990 y 1999 fueron los más calurosos en el siglo XX.* Tener cinco de los años más calurosos entre 1990 y 1999 es estadísticamente significativo. Sólo por azar, cualquier año en un siglo tendría una posibilidad de 5 en 100, o 1 en 20 de ser uno de los cinco años más calurosos. Al tener cinco de esos años en la misma década es muy poco probable que sea sólo por azar. Esta significancia estadística sugiere que el mundo podría estarse calentando.
- c. *El equipo con el peor registro de ganados-perdidos en el básquetbol gana un juego contra el campeón de la liga.* Este triunfo no es estadísticamente significativo ya que, aunque esperamos que un equipo con un pobre registro de ganados-perdidos pierda la mayoría de sus juegos, también esperamos que en ocasiones gane, incluso contra los campeones de la liga.

De la muestra a la población

Revisemos la idea de significancia estadística en una encuesta de opinión. Suponga que en una encuesta de 1 000 personas elegidas al azar, 51% apoya al presidente. Una semana después, en otra encuesta con una muestra diferente elegida al azar de 1 000 personas, sólo 49% lo apoyan. ¿Debe concluir que la opinión de los estadounidenses cambió durante la semana ocurrida entre las encuestas?

Quizá pueda adivinar que la respuesta es *no*. Los resultados de la encuesta son *estadísticas muestrales* (vea la sección 1.1): 51% de las personas en la primera muestra apoyan al presidente. Podemos utilizar este resultado para estimar el *parámetro de la población*, que es el porcentaje de *todos* los estadounidenses quienes apoyan al presidente. En el mejor de los casos, si la primera encuesta se realizó bien, diría que el porcentaje de estadounidenses que apoyan al presidente es *cercano* a 51%. De forma similar, el 49% que resulta en la segunda encuesta significa que el porcentaje de estadounidenses que apoyan al presidente es *cercano* a 49%. Puesto que dos estadísticas muestrales sólo difieren ligeramente (51% frente a 49%), es muy probable que el porcentaje real que apoya al presidente no cambiara. En lugar de eso, las dos encuestas reflejan diferencias esperadas y razonables entre las dos muestras.

En contraste, suponga que la primera encuesta halló que 75% de la muestra apoyaba al presidente y la segunda encuesta determinó que sólo 30% lo apoyaba. Suponiendo que ambas encuestas se realizaron cuidadosamente, es poco probable que los dos grupos de 1 000 personas seleccionadas de manera aleatoria pudiesen diferir tanto sólo por el azar. En este caso buscaríamos otra explicación. Quizá las opiniones de los estadounidenses cambiaron en la semana entre las encuestas.

En términos de significancia estadística, el cambio de 51% a 49% en el primer conjunto de encuestas *no* es estadísticamente significativo, ya que podemos atribuir, razonablemente, este cambio a variaciones aleatorias entre las dos muestras. Sin embargo, en el segundo conjunto de encuestas, el cambio de 75% a 30% es estadísticamente significativo, ya que es poco probable que haya ocurrido por el azar.

NOTA TÉCNICA

La diferencia entre 49% y 51% no es estadísticamente significativa para encuestas típicas, pero puede serlo para una encuesta que incluya una muestra muy grande. En general, cualquier diferencia puede ser significativa si el tamaño de la muestra es suficientemente grande.

EJEMPLO 2 Significancia estadística en experimentos

Un investigador realiza un experimento doble ciego que prueba si una nueva fórmula de hierbas es efectiva en la prevención de resfriados. Durante un periodo de tres meses, las 100 personas elegidas aleatoriamente en un grupo de tratamiento toman la fórmula de hierbas, mientras que las 100 personas elegidas aleatoriamente en un grupo de control toman un placebo. Los resultados muestran que 30 personas en el grupo de tratamiento se resfrían, comparadas con 32 personas en el grupo de control. ¿Puede concluir que la fórmula de hierbas es efectiva en la prevención de resfriados?

Solución El que una persona se resfríe durante cualquier periodo de tres meses depende de muchos factores impredecibles. Por tanto, no debemos esperar que el número de personas con resfriados en cualesquiera dos grupos de 100 personas sea exactamente el mismo. En este caso, la diferencia entre 30 personas resfriadas en el grupo de tratamiento y 30 personas con resfriado en el grupo de control parece ser suficientemente pequeña para que sea explicada por el azar. Así que la diferencia no es estadísticamente significativa, y no debemos concluir que el tratamiento sea efectivo.

Cuantificación de la significancia estadística

En el ejemplo 2 dijimos que la diferencia entre 30 resfriados en el grupo de tratamiento y 32 resfriados en el grupo de control no era estadísticamente significativa. Esta conclusión fue casi obvia a consecuencia de que la diferencia fue muy pequeña. Pero suponga que 24 personas en el grupo de tratamiento tuvieron resfriado, comparado con el 32 en el grupo de control. ¿La diferencia entre 24 y 32 será suficientemente grande para tener significado estadístico? La definición de significancia estadística que hemos utilizado hasta ahora es demasiado vaga para responder tal pregunta. Necesitamos una forma de cuantificar la idea de significancia estadística.

En general, determinamos la significancia estadística usando probabilidad para cuantificar la verosimilitud de que el resultado pueda haber ocurrido por el azar. Por tanto, hacemos la siguiente pregunta: *¿la probabilidad de que la diferencia observada ocurrió por el azar es menor o igual a 0.05 (o 1 en 20)?* Si la respuesta es *sí* (la probabilidad es menor o igual a 0.05), entonces decimos que la diferencia es *estadísticamente significativa al nivel 0.05*. Si la respuesta es *no*, la diferencia observada es razonablemente probable que haya ocurrido por el azar, por lo que decimos que no es estadísticamente significativa.

La probabilidad 0.05 es un tanto arbitraria, pero es una cifra que los estadísticos utilizan con frecuencia. Sin embargo, en ocasiones se utilizan otras probabilidades, tal como 0.1 o 0.01.

*Aquel que no deja nada al azar
hará pocas cosas mal, pero hará
muy pocas cosas.*

—George Savile Halifax

Estadísticamente significativa al nivel 0.01 es más fuerte que significativa al nivel 0.05, que es más fuerte que significativa al nivel 0.1.

Cuantificación de la significancia estadística

- Si la probabilidad de que una diferencia observada ocurra por el azar es 0.05 (o 1 en 20) o menos, la diferencia es estadísticamente significativa al nivel 0.05.
- Si la probabilidad de que una diferencia observada ocurra por el azar es 0.01 (o 1 en 100) o menos, la diferencia es estadísticamente significativa al nivel 0.01.

Quizá pueda ver que se debe tener cautela cuando se trabaja con significancia estadística. Esperaríamos que aproximadamente 1 en 20 ensayos dé un resultado estadísticamente significativo al nivel 0.05, aunque los resultados ocurran por el azar. Así, la significancia estadística al nivel 0.05 —o a casi cualquier nivel, para esa cuestión— *no garantiza* que esté presente un efecto o una diferencia importante.

UN MOMENTO DE REFLEXIÓN

Suponga que un experimento determina que las personas que toman un nuevo remedio de hierbas se resfrían menos que personas que toman un placebo, y los resultados son estadísticamente significativos al nivel 0.01. ¿El experimento ha *probado* que el remedio de hierbas funciona? Explique.

EJEMPLO 3 Significancia de la vacuna contra la polio

En la prueba de la vacuna contra la polio de Salk (vea la sección 1.1), 33 de los 200 000 niños en el grupo de tratamiento adquirieron parálisis por polio, mientras que 115 de los 200 000 en el grupo de control adquirieron parálisis por polio. Los cálculos muestran que la probabilidad de que esta diferencia entre los grupos ocurra por azar es menor que 0.01. Describa las implicaciones de este resultado.

Solución Los resultados de la prueba de la vacuna contra la polio son estadísticamente significativos al nivel 0.01, lo cual significa que hay una probabilidad 0.01 (o menos) de que la diferencia entre los grupos de control y de tratamiento ocurrió por el azar. Por tanto, podemos tener confianza que la vacuna en realidad fue la responsable del menor número de casos de polio en el grupo de tratamiento. (De hecho, la probabilidad de que los resultados de Salk ocurriesen por el azar es *mucho* menor que 0.01, por lo que los investigadores estuvieron muy convencidos de que la vacuna funcionaba; como analizaremos en el capítulo 9, esta probabilidad se denomina un valor P).

Sección 6.1 Ejercicios

Alfabetización estadística y pensamiento crítico

1. **Significancia estadística.** Un experimento prueba un método de selección de género con la intención de incrementar la posibilidad de que un bebé sea mujer, 20 parejas obtuvieron 11 niñas y 9 niños. Un representante de la compañía afirma que esto evidencia que el método es efectivo, ya que el porcentaje de niñas excede el 50% que es normalmente esperado. ¿Coincide con esta afirmación? ¿Por qué sí o por qué no?
2. **Significancia estadística.** ¿El término *significancia estadística* se refiere a resultados que son significativos en el sentido que tienen gran importancia? Explique.

3. **Significancia estadística.** Si un resultado particular es estadísticamente significativo al nivel 0.05, ¿también debe ser estadísticamente significativo al nivel 0.01? ¿Por qué sí o por qué no?
4. **Significancia estadística.** Si un resultado particular es estadísticamente significativo al nivel 0.01, ¿también debe ser estadísticamente significativo al nivel 0.05? ¿Por qué sí o por qué no?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Manejo en estado de ebriedad.** La tasa de accidentes mortales por conducir en estado de ebriedad tiene significancia estadística, ya que afecta seriamente a muchas personas.
6. **Significancia estadística.** Un experimento prueba un método para la selección de género, 100 parejas dan a luz a 80 niñas y 20 niños. Puesto que esos resultados son estadísticamente significativos, no pudieron haber ocurrido por azar.
7. **Selección de género.** En una prueba de una técnica de selección de género, de los 100 bebés nacidos al menos 80 son niñas. Puesto que existe alrededor de una posibilidad en mil millones de obtener al menos 80 niñas entre 100 bebés, los resultados son estadísticamente significativos.
8. **Ensayo clínico.** En un ensayo clínico de un tratamiento para reducir el dolor de espalda, la diferencia entre el grupo de tratamiento y el grupo de control (sin tratamiento) se encontró que era estadísticamente significativa. Esto quiere decir que el tratamiento eliminará con facilidad el dolor de espalda.

Conceptos y aplicaciones

Significancia subjetiva. Para cada evento en los ejercicios 9 al 16, indique si la diferencia entre lo que ocurrió y lo que habría esperado por el azar es estadísticamente significativa. Analice cualesquiera implicaciones de la significancia estadística.

9. **Lanzamiento de monedas.** En 500 lanzamientos de una moneda, observa 255 cruces.
10. **Lanzamiento de monedas.** En 500 lanzamientos de una moneda, observa 400 cruces.
11. **Tiros de un dado.** En 60 tiros de un dado con seis caras, el resultado 3 nunca aparece.
12. **Giro de una ruleta.** Una ruleta tiene 38 sectores numerados 0, 00, 1, 2, 3, . . . , 36. En cinco giros de una ruleta ocurrió el resultado 5 de manera consecutiva.
13. **Encuesta.** Al realizar una encuesta de adultos en Estados Unidos, un encuestador afirma que él seleccionó de manera aleatoria a 20 sujetos y todos ellos fueron mujeres.
14. **Clase de estadística.** Una instructora caminó en su primera clase de estadística y encontró que todos los estudiantes eran mujeres con al menos 6 pies de estatura.
15. **Composición de un jurado.** Para un caso sobre el cargo de no pagar la manutención de los hijos, el jurado consistió de exactamente 6 hombres y 6 mujeres.
16. **Ensayo clínico.** En un ensayo clínico de una droga para tratar alergias, 5 de los 80 sujetos en el grupo de tratamiento experimentaron dolor de cabeza, y 8 de los 160 sujetos en el grupo de control experimentaron dolor de cabeza.
17. **Pruebas de gasolina.** Treinta automóviles idénticos se seleccionan para una prueba de gasolina. La mitad de ellos se llenan con gasolina regular, la otra mitad se llena con una nueva gasolina experimental. Los automóviles en el primer grupo promedian 29.3 millas por galón, mientras que los del segundo grupo promedian 35.5 millas por galón. Analice si esta diferencia parece estadísticamente significativa.
18. **Tratamientos del síndrome del túnel carpiano.** Se realiza un experimento para determinar si hay diferencia entre las tasas de éxito en el tratamiento del síndrome del túnel carpiano con cirugía y con entablillado. La tasa de éxito para 73 pacientes tratados con cirugía fue 92% y la tasa de éxito para 83 pacientes tratados con entablillado fue 72% (con base en datos de “Splinting vs. Surgery in the Treatment of Carpal Tunnel Syndrome” de Gerritsen, *et al.*, *Journal of the American Medical Association*, volumen 288, número 10). Analice si esta diferencia parece ser estadísticamente significativa.
19. **Selección de género.** El Instituto de Genética e IVF realizó un ensayo clínico de su método para la selección de género. Cuando se escribió este libro, 325 bebés habían nacido de padres usando el método XSORT, para aumentar la probabilidad de concebir a una niña, y 295 de esos bebés fueron niñas. Analice si estos resultados parecen ser estadísticamente significativos.
20. **Mosquitero y malaria.** En un ensayo aleatorizado y controlado en Kenia, se probaron mosquiteros tratados con insecticida como una manera de reducir la malaria. Entre 343 niños que utilizaron mosquiteros, 15 desarrollaron malaria. Entre 294 infantes que no los utilizaron, 27 adquirieron la malaria (con base en información de “Sustainability of Reductions in Malaria Transmission and Infant Mortality in Western Kenya with Use of Insecticide-Treated Bednets” de Lindblade *et al.*, *Journal of the American Medical Association*, volumen 291, número 21). Suponiendo que las redes de cama no tienen efecto, hay una probabilidad 0.015 de obtener estos resultados debido al azar. ¿Los resultados parecen ser estadísticamente significativos? ¿Parece que los mosquiteros son efectivos?
21. **Temperatura del cuerpo humano.** En un estudio realizado por investigadores en la Universidad de Maryland, se midió las temperaturas corporales; la media para la muestra fue 98.20°F. El valor aceptado para la temperatura del cuerpo humano es 98.60°F. La diferencia entre la media muestral y el valor aceptado es significativa al nivel 0.05.
 - a. Analice el significado del nivel de significancia en este caso.
 - b. Si suponemos que la temperatura media del cuerpo en realidad es 98.6°F, la probabilidad de obtener una muestra con una media de 98.20°F o menos es 0.000000001. Interprete este valor de probabilidad.
22. **Cinturones de seguridad y niños.** En un estudio de niños lesionados en choques automovilísticos (*American Journal of Public Health*, volumen 82, número 3), aquellos que usaron cinturones de seguridad tuvieron una estancia media de 0.83 días en una unidad de cuidados intensivos. Los que no utilizaron cinturón de seguridad tuvieron una media de estancia de 1.39 días. La diferencia en medias entre los dos

grupos es significativa al nivel 0.0001. Interprete este resultado.

23. Preparación para el SAT. Un estudio de 75 estudiantes que presentaron un curso de preparación para el SAT (*American Research Journal*, vol. 19, núm. 3) concluyó que la mejora en la media en el SAT fue de 0.6 puntos. Si suponemos que el curso de preparación no tiene efecto, la probabilidad de obtener una mejora en la media de 0.6 puntos debido al azar es 0.08. Analice si este curso de preparación tiene como resultado una mejora estadísticamente significativa.

24. Peso por edad. Una encuesta nacional de salud determinó que el peso medio de una muestra de 804 hombres en edades entre 25 y 34 años fue de 176 libras, mientras que el peso medio de una muestra de 1 657 hombres en edades de 65 a 74 años fue de 164 libras. La diferencia es significativa al nivel 0.01. Interprete este resultado.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 6 en www.aw.com/bbt.

25. Significancia en estadísticas vitales. Visite un sitio web que tenga estadísticas vitales (por ejemplo, la Oficina de Censo de Estados Unidos o el Centro Nacional de Estadísticas de Salud). Seleccione una pregunta como la siguiente:

- ¿Son significativas las diferencias en el número de nacimientos entre los meses?
- ¿Son significativas las diferencias en el número de muertes naturales entre días de la semana?
- ¿Son significativas las diferencias en tasas de mortalidad infantil entre estados seleccionados?
- ¿Son significativas las diferencias en incidencias de una enfermedad particular entre estados seleccionados?

- ¿Son significativas las diferencias entre las tasas de matrimonio en varios estados?

Recolecte la información relevante y determine de manera subjetiva si considera que las diferencias observadas son significativas, esto es, explique si podrían ocurrir por efecto del azar o proporcione alguna explicación alterna.

26. Longitudes de ríos. Usando un almanaque o internet determine las longitudes de los principales ríos del mundo. Construya una lista sólo de los dígitos principales (el de más a la izquierda). ¿Algún dígito particular aparece con más frecuencia que los otros? ¿Ese dígito aparece significativamente más que los otros? Explique.

EN LAS NOTICIAS

27. Significancia estadística. Encuentre una noticia reciente del periódico sobre una investigación estadística en la que se utilice la idea de significancia estadística. Escriba un resumen de una página del estudio y el resultado que se considera estadísticamente significativo. También incluya una breve discusión acerca de su credibilidad del resultado, dada su significancia estadística.

28. ¿Experimento significativo? Encuentre una noticia reciente acerca de una investigación estadística que utilice un experimento para determinar si algún tratamiento nuevo fue efectivo. Con base en la información disponible, analice brevemente lo que puede concluir sobre la significancia estadística de los resultados. Dada esta significancia (o carencia de ella) ¿considera que el nuevo tratamiento es útil? Explique.

29. Significancia estadística personal. Describa un incidente en su vida que no cumpla con su expectativa, desafía las posibilidades o parece improbable que haya ocurrido por el azar. ¿Llamaría a este incidente estadísticamente significativo? ¿A qué atribuye el evento?

6.2 Fundamentos de probabilidad

En la sección 6.1 vimos que las ideas de probabilidad son fundamentales en estadística, en parte por el concepto de significancia estadística. Regresaremos al tema de significancia estadística en los capítulos 7 al 10. Primero necesitamos explorar unas cuantas ideas de probabilidad y sus aplicaciones en la vida diaria.

Iniciamos considerando un lanzamiento de dos monedas. La figura 6.1 muestra que existen cuatro formas diferentes en que podrían caer las monedas, decimos que cada una es un **resultado** diferente del lanzamiento de las monedas. Los resultados son los únicos posibles del lanzamiento de monedas. Pero suponga que estamos interesados sólo en el número de caras. Puesto que los dos resultados centrales en la figura 6.1 cada uno tienen una cara, decimos que estos dos resultados representan el mismo evento. Un **evento** describe uno o más resultados posibles que tienen la misma propiedad de interés, en este caso, el mismo número de caras.

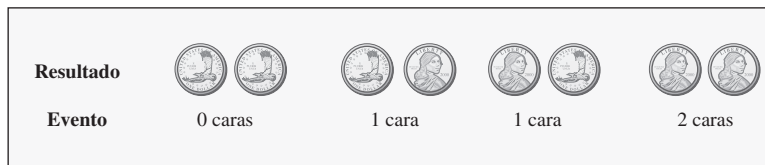


Figura 6.1 Los cuatro posibles resultados en el lanzamiento de dos monedas. Ambos resultados de en medio representan el mismo *evento* de 1 cara.

La figura 6.1 muestra que hay cuatro posibles resultados para el lanzamiento de dos monedas, pero sólo tres posibles eventos: 0 caras, 1 cara y 2 caras.

Definiciones

Resultados son las consecuencias únicas posibles de observaciones o experimentos.

Un **evento** es una colección de uno o más resultados que comparte una propiedad de interés.

En forma matemática expresamos probabilidades como números entre 0 y 1: la probabilidad de que una moneda caiga cara es un medio, o 0.5, si un evento es imposible, le asignamos una probabilidad 0. Por ejemplo, la probabilidad de encontrar a un soltero casado es 0. En el otro extremo, a un evento que es seguro que ocurra se le asigna una probabilidad 1. Por ejemplo, de acuerdo al viejo dicho de Benjamín Franklin, la probabilidad de morir y de pagar impuestos es 1. Escribimos $P(\text{evento})$ para dar a entender la probabilidad de un evento. Con frecuencia denotamos a los eventos por medio de letras o símbolos. Por ejemplo, si usamos H para representar el evento de cara en el lanzamiento de una moneda, escribimos $P(H) = 0.5$.

Cómo expresar probabilidades

La probabilidad de un evento, expresado como $P(\text{evento})$, siempre está entre 0 y 1, incluyendo a ambos. Una probabilidad 0 significa que el evento es imposible y una probabilidad 1 significa que el evento es seguro.

La figura 6.2 muestra la escala de valores de probabilidad, junto con expresiones de uso común de posibilidad. Es útil desarrollar el sentido para ver que un valor de probabilidad tal como 0.95 indica que un evento es muy probable que ocurra, pero no es seguro. Debe ocurrir en alrededor de 95 de 100 veces. En contraste, un valor de probabilidad 0.01 describe un evento que es muy poco probable que ocurra. El evento es posible, pero sólo ocurrirá una de cada 100 veces.

Con estas definiciones y notación básicas, ya estamos preparados para estudiar las tres técnicas básicas para determinar probabilidades. Llevan por nombres *método teórico*, *método de frecuencia relativa* y *método subjetivo*. Exploraremos cada uno de ellos.

Probabilidades teóricas

Cuando decimos que la probabilidad de obtener cara en el lanzamiento de una moneda es $1/2$, estamos *suponiendo* que la moneda es justa y es igualmente probable que caiga en cara o en cruz. En esencia, la probabilidad está basada en una teoría de cómo se comporta la moneda, de modo que decimos que la probabilidad $1/2$ proviene del método teórico. Como otro ejemplo, considere el tiro de un dado. Puesto que hay seis resultados igualmente probables (figura 6.3), la probabilidad teórica para cada resultado es $1/6$.

Mientras todos los resultados sean igualmente probables, podemos utilizar el procedimiento siguiente para calcular probabilidades teóricas.

En este mundo nada es seguro, excepto la muerte y los impuestos.

—Benjamín Franklin

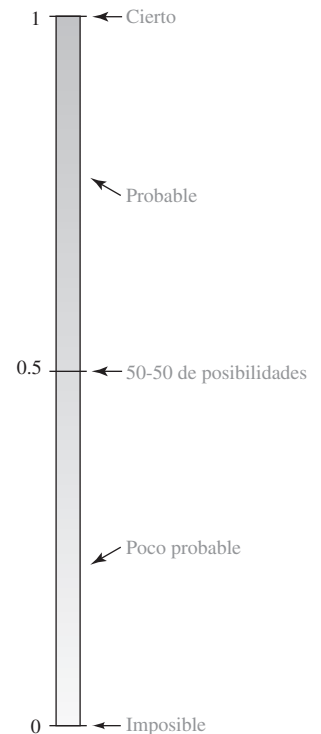


Figura 6.2 La escala muestra varios grados de certidumbre cuando se expresan como probabilidades.

A propósito...

Un análisis sofisticado del matemático Persi Diaconis ha mostrado que revolver siete veces un mazo nuevo de cartas es necesario para darle un orden en el que todos los arreglos de cartas sean igualmente probables.

A propósito...

Desde el año 3600 a.n.e, huesos redondeados, denominados *astragali* fueron usados de forma muy parecida a dados para juegos de azar en el Oriente Medio. Nuestros dados cúbicos aparecieron alrededor de 2000 a.n.e, en Egipto y en China. Los juegos de cartas fueron inventados en China en el siglo x y llevados a Europa en el siglo xiv.



A propósito...

Los días de nacimientos no se distribuyen aleatoriamente a lo largo del año. Además de las diferentes duraciones de los meses, enero tiene la tasa de nacimientos más baja, y junio y julio tienen las tasas de nacimiento más altas.

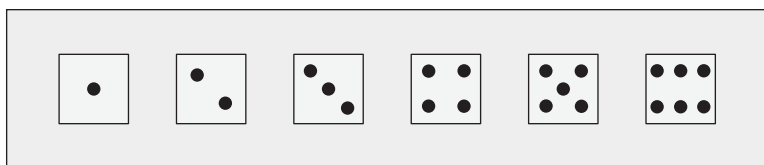


Figura 6.3 Los seis resultados posibles para el tiro de un dado con seis caras.

Método teórico para resultados igualmente probables

Paso 1. Cuente el número total de resultados posibles.

Paso 2. Entre todos los resultados posibles, cuente el número de formas en que el evento de interés, A , puede ocurrir.

Paso 3. Determine la probabilidad, $P(A)$, como

$$P(A) = \frac{\text{número de formas en que puede ocurrir } A}{\text{número total de resultados}}$$

EJEMPLO 1 Adivinación de cumpleaños

Suponga que selecciona una persona al azar entre un grupo numeroso en una conferencia. ¿Cuál es la probabilidad de que la persona elegida haya nacido en julio? Suponga que hay 365 días en un año.

Solución Si suponemos que todos los días de nacimiento son igualmente probables, podemos utilizar los tres pasos del método teórico.

Paso 1. Cada día de nacimiento representa un resultado, por lo que hay 365 resultados posibles.

Paso 2. Julio tiene 31 días, por lo que 31 de los 365 resultados posibles representan el evento de un nacimiento en julio.

Paso 3. La probabilidad de que una persona elegida al azar haya nacido en julio es

$$P(\text{nacimiento en julio}) = \frac{31}{365} \approx 0.0849$$

que es un poco más de 1 en 12.

Conteo de resultados

Suponga que lanzamos dos monedas y queremos contar el número total de resultados. El lanzamiento de la primera moneda tiene dos resultados posibles cara (H) y cruz (T). El lanzamiento de la segunda moneda también tiene dos posibles resultados. Los dos resultados de la primera moneda pueden ocurrir con cualquiera de los dos resultados de la segunda moneda. Por lo que el número total de resultados para las dos monedas es $2 \times 2 = 4$, son HH, HT, TH y TT, como se muestra en el diagrama de árbol de la figura 6.4a.

UN MOMENTO DE REFLEXIÓN

Explique por qué los resultados del lanzamiento de una moneda dos veces en sucesión son los mismos que los del lanzamiento de dos monedas al mismo tiempo.

Ahora podemos ampliar esto. Si lanzamos tres monedas, tenemos un total de $2 \times 2 \times 2 = 8$ posibles resultados: HHH, HHT, HTH, HTT, THH, THT, TTH y TTT, como se muestra en la figura 6.4b. Esta idea es la base para la regla de conteo siguiente.

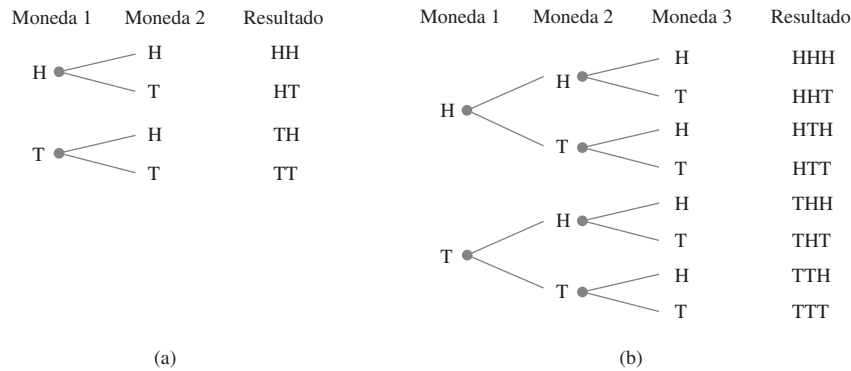


Figura 6.4 Diagramas de árbol que muestran los resultados de lanzar (a) dos y (b) tres monedas.

Conteo de resultados

Suponga que un proceso A tiene a posibles resultados y el proceso B tiene b resultados posibles. Si los resultados del proceso no se afectan entre ellos, el número de resultados diferentes para los dos procesos combinados es $a \times b$. Esa idea se amplía a cualquier número de procesos. Por ejemplo, si un tercer proceso C tiene c resultados posibles, el número de resultados posibles de los tres procesos combinados es $a \times b \times c$.

EJEMPLO 2 Algunos conteos

- ¿Cuántos resultados posibles hay si tira un dado y lanza una moneda?
- ¿Cuál es la probabilidad de obtener dos unos (ojos de serpiente) al tirar dos dados?

Solución

- El primer proceso, tirar un dado, tiene seis resultados (1, 2, 3, 4, 5, 6). El segundo proceso de lanzar una moneda, tiene dos resultados (H, T). Por tanto existen $6 \times 2 = 12$ resultados para los dos procesos juntos (1H, 1T, 2H, 2T, ..., 6H, 6T).
- El tiro de un solo dado tiene seis resultados igualmente probables. Por tanto, cuando se tiran dos dados, existen $6 \times 6 = 36$ resultados diferentes. De estos 36 resultados, sólo uno es el evento de interés (dos unos). Por lo que la probabilidad de obtener dos unos es

$$P(\text{dos unos}) = \frac{\text{número de maneras de que puedan ocurrir dos unos}}{\text{número total de resultados}} = \frac{1}{36} = 0.0278$$

EJEMPLO 3 Conteo de hijos

¿Cuál es la probabilidad de que, en una familia con tres hijos seleccionada aleatoriamente, el mayor sea hombre, la segunda mujer y la más joven mujer? Suponga que hombres y mujeres son igualmente probables.

Solución Existen dos posibles resultados para cada nacimiento: hombre o mujer. Para una familia con tres hijos, el número total de resultados posibles (órdenes de nacimiento) es $2 \times 2 \times 2 = 8$ (HHH, HHM, HMH, HMM, MHH, MHM, MMH, MMM). La pregunta se hace acerca de un orden de nacimientos particular (hombre-mujer-mujer), así que esto es 1 de 8 posibles resultados. Por tanto, este orden de nacimientos tiene una probabilidad $1/8 = 0.125$.

A propósito...

Los nacimientos de hombres y mujeres *no* son igualmente probables. De manera natural suceden aproximadamente 105 nacimientos de varones por cada 100 nacimientos de mujeres. Sin embargo, la tasa de mortalidad masculina es más alta que la tasa de mortalidad femenina, por lo que las mujeres adultas superan en número a los hombres adultos.



UN MOMENTO DE REFLEXIÓN

¿Cuántas familias diferentes de cuatro hijos son posibles, si el orden de nacimiento se toma en cuenta? ¿Cuál es la probabilidad de que una pareja con cuatro descendientes, sean mujeres?

Probabilidades con frecuencias relativas

La segunda manera de determinar probabilidades es *aproximar* la probabilidad un evento A mediante la realización de muchas observaciones y contar el número de veces que ocurre el evento A . Este enfoque se denomina **método de la frecuencia relativa** (o **empírico**). Por ejemplo, si observamos que llueve un promedio de 100 días al año, podríamos decir que la probabilidad de que llueva en un día seleccionado aleatoriamente es $100/365$. Podemos utilizar una regla general para este método.

Método de la frecuencia relativa

Paso 1. Repita u observe un proceso muchas veces y cuente el número de veces que el evento de interés, A , ocurre.

Paso 2. Estime $P(A)$ mediante

$$P(A) = \frac{\text{número de veces que ocurrió } A}{\text{número total de observaciones}}$$

EJEMPLO 4 Inundación cada 500 años

Los registros geológicos indican que un río ha crecido por arriba de un nivel particular de gran inundación cuatro veces en los últimos 2000 años. ¿Cuál es la probabilidad de frecuencia relativa que el río sobrepase el nivel de gran inundación el año siguiente?

Solución Con base en la información, la probabilidad de que el río crezca por arriba del nivel de gran inundación en cualquier año particular es

$$\frac{\text{número de años con inundación}}{\text{número total de años}} = \frac{4}{2000} = \frac{1}{500}$$

Puesto que este tipo de una inundación ocurre en promedio una vez cada 500 años, se denomina “inundación de 500 años”. La probabilidad de tener una inundación de esta magnitud en cualquier año dado es $1/500$, o 0.002 .

Probabilidades subjetivas

El tercer método para la determinación de probabilidades es estimar una **probabilidad subjetiva** usando la experiencia o la intuición. Por ejemplo, podría hacer una estimación subjetiva de la probabilidad de que un amigo se case el año siguiente o de la probabilidad de que una buena calificación en estadística le ayude a obtener el empleo que quiere.

A propósito...

Los métodos teóricos también se denominan métodos *a priori*. Las palabras *a priori* son del latín para “antes del hecho” “antes de la experiencia”.

Tres enfoques para determinar la probabilidad

Una **probabilidad teórica** está basada en la suposición de que todos los resultados son igualmente probables. Para determinarla se divide el número de maneras en que un evento puede ocurrir entre el número total de resultados posibles.

Una **probabilidad de frecuencia relativa** se basa en observaciones o experimentos. Es la frecuencia relativa del evento de interés.

Una **probabilidad subjetiva** es una estimación basada en la experiencia o intuición.

EJEMPLO 5 ¿Cuál método?

Identifique el método que resultó en los enunciados siguientes.

- La probabilidad de que usted se casará el año próximo es cero.
- Con base en información del gobierno, la posibilidad de morir en un accidente automovilístico es 1 en 7 000 (por año).
- La posibilidad de obtener un 7 con un dado de doce caras es $1/12$.

Solución

- Ésta es una probabilidad subjetiva, ya que tiene como base un sentimiento en el momento actual.
- Ésta es una probabilidad de frecuencia relativa, ya que tiene como base datos en accidentes automovilísticos pasados.
- Ésta es una probabilidad teórica, porque está basada en la hipótesis de que un dado de doce caras es igualmente probable que caiga en cualquiera de las doce caras.

A propósito...

Otro enfoque para determinar probabilidades, denominado *método Monte Carlo*, utiliza simulaciones con computadora. Esta técnica, en esencia, determina probabilidades de frecuencia relativa; en este caso, las observaciones se hacen con la computadora.

EJEMPLO 6 Probabilidades de huracán

La figura 6.5 muestra un mapa publicado cuando el huracán Floyd se aproximó a la costa sureste de Estados Unidos. El mapa muestra “probabilidades de choque” en tres regiones diferentes a lo largo de la trayectoria del huracán. Interprete el término “probabilidad de choque” y analice el valor de estas probabilidades y cómo se estimaron.

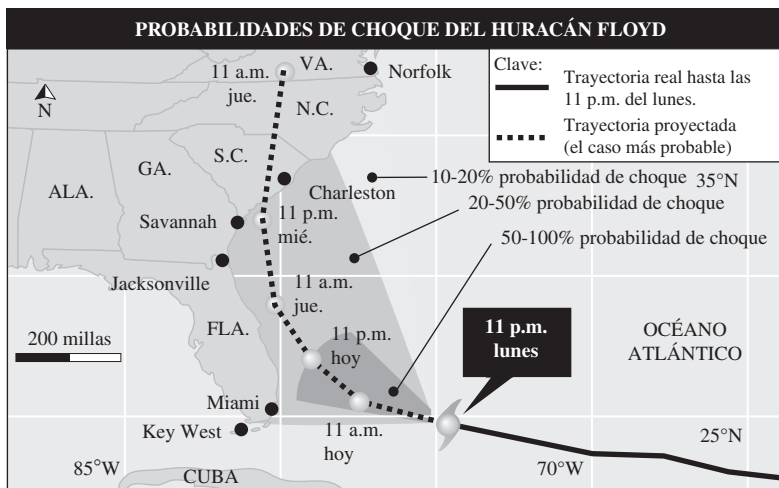


Figura 6.5 Este mapa muestra la trayectoria real del huracán Floyd actualizada al momento en que el mapa fue realizado (“11 p.m. del lunes”). Entonces, muestra las probabilidades de varias trayectorias alternas, con la trayectoria más probable en línea gruesa discontinua. Fuente: Virginia Pilot.

Solución Las probabilidades de choque indican cuán probable es que el ojo del huracán pase sobre una localidad particular. Por ejemplo, para una persona en la región de probabilidad de choque de 20-50%, existe una probabilidad de 0.2 a 0.5 de ser golpeada por el huracán. El conocimiento de estas probabilidades ayuda a la gente a tomar decisiones acerca de preparativos y posibles evacuaciones. Las probabilidades se calculan con modelos de cómputo que predicen el comportamiento del huracán con base en los patrones actuales del clima en la región del huracán, la física de los huracanes y datos observados de huracanes anteriores.

Probabilidad de que un evento *no* ocurra

Suponga que estamos interesados en la probabilidad de que un evento o resultado particular *no* ocurra. Por ejemplo, considere la probabilidad de una respuesta errónea a una pregunta de opción múltiple con cinco respuestas posibles. La probabilidad de responder correctamente con una respuesta al azar es 1/5, por lo que la probabilidad de *no* responder correctamente es 4/5. Observe que la probabilidad de responder que una respuesta sea correcta o una respuesta sea incorrecta es 1. Podemos generalizar esta idea.

Probabilidad de que un evento *no* ocurra

Si la probabilidad de un evento *A* es $P(A)$, entonces la probabilidad de que el evento *A* no ocurra es $P(\text{no } A)$. Puesto que el evento debe ocurrir o bien no ocurrir, podemos escribir

$$P(A) + P(\text{no } A) = 1 \quad \text{o} \quad P(\text{no } A) = 1 - P(A)$$

Nota: el evento *no A* se denomina el **complemento del evento A**; la palabra “no” a veces se denota mediante una barra, por lo que $\bar{A}$ significa *no A*.

A propósito...

Un *lote* es una parte o asignación de una persona. El término viene a ser utilizado para cualquier objeto o señal que identifica a una persona en una selección al azar. El reparto de lotes fue utilizado en las guerras troyanas para determinar quién lanzaría la primera lanza y en el Antiguo Testamento para dividir tierras conquistadas.

EJEMPLO 7

En un estudio de sistemas de verificación mediante escáner, muestras de compras se utilizaron para comparar los precios escaneados con los precios anunciados. La tabla 6.1 resume los resultados para una muestra de 819 artículos. Con base en estos datos, ¿cuál es la probabilidad de que un artículo marcado con precio regular tenga un error de escaneo? ¿Cuál es la probabilidad de que un artículo anunciado con precio especial tenga un error de escaneo?

Tabla 6.1 Precisión de escaneo		
	Artículos con precio regular	Artículos anunciados con precio especial
Subvaluado	20	7
Sobrevaluado	15	29
Precio correcto	384	364

Fuente: Ronald Goodstein, “UPC Scanner Pricing Systems: Are They Accurate?”, *Journal of Marketing*, vol. 58.

Solución Podemos representar con *R* un artículo de precio regular que sea escaneado correctamente. Ya que 384 de 419 artículos con precio regular se escanearon correctamente,

$$P(R) = \frac{384}{419} = 0.916$$

El caso de un error de análisis es el complemento del evento de una correcta exploración, por lo que la probabilidad de que un artículo sometido a precio regular es de un error de análisis (ya sea cobrado de menos o sobrecarga) es

$$P(\text{no } R) = 1 - 0.916 = 0.084$$

Suponga que *A* representa a un artículo con precio especial que se escaneó correctamente. De los 400 artículos especiales anunciados en la muestra, 364 se escanearon correctamente. Por tanto,

$$P(A) = \frac{364}{400} = 0.910$$

La probabilidad de que un artículo especial anunciado sea escaneado de manera incorrecta es

$$P(\text{no } A) = 1 - 0.910 = 0.090$$

Las tasas de errores casi son iguales. Sin embargo, observe en la tabla 6.1 que la mayoría de los errores cometidos con los artículos especiales anunciados no son en favor del cliente. Con artículos con precios regulares, más de la mitad de los errores son a favor del cliente.

Distribuciones de probabilidad

En los capítulos 3 al 5 trabajamos con distribuciones de frecuencia y frecuencia relativa, por ejemplo, distribuciones de edad o de ingreso. Una de las ideas más fundamentales en probabilidad y estadística es la *distribución de probabilidad*. Como el nombre sugiere, una distribución de probabilidad es una distribución en que la variable de interés se asocia con una probabilidad.

Suponga que lanza dos monedas de manera simultánea. Los resultados son las diferentes combinaciones de una cara y una cruz en las dos monedas. Puesto que *cada* moneda puede caer en dos formas posibles (cara o cruz), las *dos* monedas pueden caer de $2 \times 2 = 4$ formas diferentes. La tabla 6.2 tiene un renglón para cada uno de estos cuatro resultados.

Tabla 6.2 Resultados del lanzamiento de dos monedas			
Moneda 1	Moneda 2	Resultado	Probabilidad
H	H	HH	1/4
H	T	HT	1/4
T	H	TH	1/4
T	T	TT	1/4

Observe que los cuatro resultados representan sólo tres eventos diferentes: 2 caras (HH), 2 cruces (TT) y una cara y una cruz (HT o TH). Por tanto, la probabilidad de dos caras es $P(HH) = 1/4 = 0.25$; la probabilidad de dos cruces es $P(TT) = 1/4 = 0.25$; y la probabilidad de una cara y una cruz es $P(H \text{ y } T) = 2/4 = 0.50$. Estas probabilidades dan como resultado una **distribución de probabilidad** que puede mostrarse como una tabla (tabla 6.3) o un histograma (figura 6.6). Observe que la suma de todas las probabilidades debe ser 1 (ya que debe ocurrir alguno de los resultados posibles).

Tabla 6.3 Lanzamiento de dos monedas	
Resultado	Probabilidad
2 caras, 0 cruces	0.25
1 cara, 1 cruz	0.5
0 caras, 2 cruces	0.25
Total	1

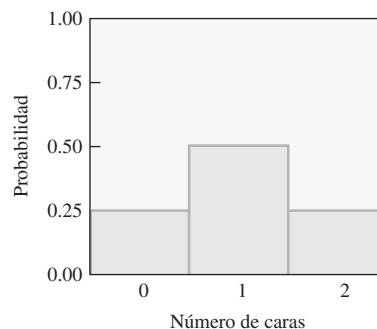


Figura 6.6 El histograma muestra la distribución de probabilidad para los resultados del lanzamiento de dos monedas.

A propósito...

Otra manera común de expresar verosimilitud es mediante los momios o posibilidades. Las posibilidades *contra* un evento es la razón de la probabilidad de que el evento no ocurra a la probabilidad de que el evento sí ocurra. Por ejemplo, las posibilidades contra tirar un 6 con un dado son $(5/6) / (1/6)$, o 5 a 1. Las posibilidades usadas en juegos se denominan *posibilidades (momios) de pago*, expresan su ganancia neta en una apuesta ganadora. Por ejemplo, suponga que los momios de pago en un caballo particular en una carrera de caballos son 3 a 1. Esto significa que por cada \$1 que usted apueste a este caballo, ganará \$3, si el caballo gana (y obtendrá \$1 de los originales de vuelta).



Construcción de una distribución de probabilidad

Una **distribución de probabilidad** representa las probabilidades de todos los eventos posibles. Haga lo siguiente para construir una exhibición de una distribución de probabilidad.

Paso 1. Liste todos los *resultados* posibles. Si le ayuda utilice una tabla o una figura.

Paso 2. Identifique los resultados posibles que representan al mismo *evento*. Determine la probabilidad cada evento.

Paso 3. Construya una tabla en la cual una columna liste cada evento y otra columna liste cada probabilidad. La suma de todas las probabilidades debe ser 1.

EJEMPLO 8
Lanzamiento de tres monedas

Construya una distribución de probabilidad para el número de caras que ocurren cuando se lanzan tres monedas de manera simultánea.

Solución Aplicamos el proceso de tres pasos.

- Paso 1. El número de resultados diferentes cuando se lanzan tres monedas es $2 \times 2 \times 2 = 8$. La figura 6.4b (página 241) muestra cómo encontramos los ocho resultados posibles, que son HHH, HHT, HTH, HTT, THH, THT, TTH, TTT.
- Paso 2. Existen cuatro eventos posibles: 0 caras, 1 cara, 2 caras y 3 caras. Observando los ocho resultados posibles, encontramos lo siguiente:

 - Sólo uno de los ocho resultados posibles representa el evento de 0 caras, por lo que la probabilidad es $1/8$.
 - Tres de los ocho resultados posibles representan el evento de 1 cara (y 2 cruces): HTT, THT y TTH. Por tanto, este evento tiene una probabilidad de $3/8$.
 - Tres de los ocho resultados posibles representan el evento de 2 caras (y 1 cruz): HHT, HTH, THH. Este evento también tiene una probabilidad de $3/8$.
 - Sólo uno de los ocho resultados representa el evento de 3 caras, por lo que su probabilidad es $1/8$.
- Paso 3. Construimos una tabla con los cuatro eventos listados en la columna izquierda y sus probabilidades en la columna derecha. La tabla 6.4 muestra el resultado.

Tabla 6.4 Lanzamiento de tres monedas	
Resultado	Probabilidad
3 caras (0 cruces)	$1/8$
2 caras (1 cruz)	$3/8$
1 cara (2 cruces)	$3/8$
0 caras (3 cruces)	$1/8$
Total	1

UN MOMENTO DE REFLEXIÓN

Cuando lanza cuatro monedas, ¿cuántos resultados diferentes son posibles? Si registra el número de caras, ¿cuántos eventos diferentes son posibles?

EJEMPLO 9
Distribución de dos dados

Construya una distribución de probabilidad para la suma de los dados cuando se tiran dos dados. Exprese la distribución como una tabla y como un histograma.

Solución Puesto que existen seis formas en que sale cada dado, existen $6 \times 6 = 36$ resultados del tiro de dos dados. Enumeramos los 36 resultados en la tabla 6.5, listando un dado en

Tabla 6.5 Resultados y sumas para el tiro de dos dados						
	1	2	3	4	5	6
1	1 + 1 = 2	1 + 2 = 3	1 + 3 = 4	1 + 4 = 5	1 + 5 = 6	1 + 6 = 7
2	2 + 1 = 3	2 + 2 = 4	2 + 3 = 5	2 + 4 = 6	2 + 5 = 7	2 + 6 = 8
3	3 + 1 = 4	3 + 2 = 5	3 + 3 = 6	3 + 4 = 7	3 + 5 = 8	3 + 6 = 9
4	4 + 1 = 5	4 + 2 = 6	4 + 3 = 7	4 + 4 = 8	4 + 5 = 9	4 + 6 = 10
5	5 + 1 = 6	5 + 2 = 7	5 + 3 = 8	5 + 4 = 9	5 + 5 = 10	5 + 6 = 11
6	6 + 1 = 7	6 + 2 = 8	6 + 3 = 9	6 + 4 = 10	6 + 5 = 11	6 + 6 = 12

los renglones y el otro en las columnas. En cada celda mostramos la suma de los números en los dos dados.

Los eventos *posibles* son las sumas de 2 a 12. Éstos son los *eventos* en este problema. Encontramos la probabilidad de cada evento contando todos los resultados para cada suma y luego dividiendo el número de resultados entre 36. Por ejemplo, los cinco resultados resaltados en la tabla tienen una suma de 8, por lo que la probabilidad de una suma de 8 es $5/36$. La tabla 6.6 muestra la distribución de probabilidad completa, y la figura 6.7 muestra la distribución como un histograma.

Tabla 6.6 Distribución de probabilidad para la suma de dos dados

Evento (suma)	2	3	4	5	6	7	8	9	10	11	12	Total
Probabilidad	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

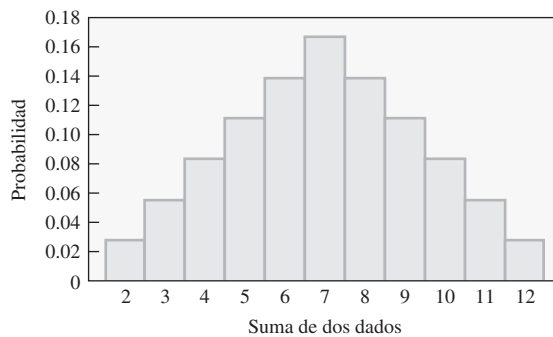


Figura 6.7 El histograma que muestra la distribución de probabilidad para la suma de dos dados.

Sección 6.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Notación.** ¿Qué representa la notación $P(A)$? ¿Qué representa la notación $P(\bar{A})$, también escrita como $P(\bar{A})$?
- Interpretación de la probabilidad.** ¿Qué queremos decir con “la probabilidad de ganar el gran premio en la lotería de Illinois es $1/20358520$ ”? ¿Ganar es tan *inusual*? ¿Por qué sí o por qué no?
- Probabilidad de lluvia.** Cuando se escribió acerca de la probabilidad de que lloviese en Boston el 4 de julio del año siguiente, un reportero indicó que la probabilidad es $1/2$, ya que llovería o no llovería. ¿Este razonamiento es correcto? ¿Por qué sí o por qué no?
- Probabilidad subjetiva.** Utilice un juicio subjetivo para estimar la probabilidad de que la siguiente vez que usted suba a un elevador se quede atascado entre los pisos.

¿Tiene sentido? Para los ejercicios 5 al 10 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los

enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Evento imposible.** Puesto que es imposible obtener un total de 14 cuando se tiran dos dados ordinarios, la probabilidad de 14 es 0.
- Colisión de automóvil.** Una compañía de seguros indica que la probabilidad de que un automóvil particular se involucre en un choque este año es 0.14 y la probabilidad de que el automóvil no se involucre en un choque este año es 0.82.
- Opuestos.** Si $P(A) = 0.25$, entonces $P(\text{no } A) = 0.75$.
- Evento cierto.** Si $P(A) = 1$, entonces el evento A definitivamente ocurre.
- Relámpago.** Jack estima que la probabilidad subjetiva de ser alcanzado por un rayo en algún momento durante el año siguiente es $1/2$.
- Relámpago.** Jill estima que la probabilidad subjetiva de ser alcanzada por un rayo en algún momento durante el año entrante es $1/1\,000\,000$.

Conceptos y aplicaciones

11. Resultados o eventos. Para cada observación descrita, indique si hay una o más de una forma para que la observación dada ocurra (en cuyo caso es un evento).

- Obtener una cara en el lanzamiento de una moneda.
- Obtener dos caras con el lanzamiento de tres monedas.
- Obtener un 6 con el tiro de un dado.
- Obtener un número par con el tiro de un dado.
- Obtener una suma de 7 con el tiro de dos dados.
- Sacar un par de reyes de un mazo regular de cartas.

12. Resultados o eventos. Para cada observación descrita, indique si hay una o más de una forma para que la observación dada ocurra (en cuyo caso es un evento).

- Obtener una cruz en el lanzamiento de una moneda.
- Obtener una cruz con el lanzamiento de tres monedas.
- Obtener al menos 5 con el tiro de dos dados.
- Obtener un número impar con el tiro de un dado.
- Sacar un jack de un mazo regular de cartas.
- Sacar tres ases de un mazo regular de cartas.

Probabilidades teóricas. Para los ejercicios del 13 al 20 utilice el método teórico con el fin de determinar la probabilidad del resultado o el evento dado. Indique cualquier suposición que necesite hacer.

- Dado.** Tirar un dado y obtener un resultado que sea menor a 3.
- Dado.** Tirar un dado y obtener un resultado que sea mayor a 6.
- Ruleta.** Obtener un resultado de 0 o 00 cuando se hace girar una ruleta. (Una ruleta tiene sectores de 0, 00, 1, 2, 3,..., 36).
- Día de nacimiento.** La determinación que el siguiente presidente de Estados Unidos haya nacido en sábado.
- Día de nacimiento.** La determinación que la siguiente persona que encuentre tenga el mismo día de nacimiento que usted. (Ignore años bisiestos).
- Nacimientos.** La determinación que el siguiente niño nacido en Alaska sea mujer.
- Nacimientos.** La determinación que el siguiente nacimiento de una pareja sea niña, dado que la pareja ya tiene dos hijos y ambos son varones.
- Prueba.** La selección aleatoria de la respuesta correcta en un examen de opción múltiple con respuestas posibles de a, b, c, d y e, una de las cuales es correcta.

Eventos opuestos. Para los ejercicios del 21 al 28 determine la probabilidad del evento opuesto dado. Indique cualesquiera suposiciones que utilice.

- Dado.** ¿Cuál es la probabilidad de tirar un dado y no obtener un resultado menor a 3?
- Dado.** ¿Cuál es la probabilidad de tirar un dado y no obtener un resultado mayor a 6?
- Ruleta.** ¿Cuál es la probabilidad de girar una ruleta y no obtener un resultado de 0 o 00? (Una ruleta tiene sectores de 0, 00, 1, 2, 3,..., 36).
- Cumpleaños.** ¿Cuál es la probabilidad de que el siguiente presidente de Estados Unidos no haya nacido en sábado?
- Básquetbol.** ¿Cuál es la probabilidad de que una tiradora de tiros libres con 55% de efectividad fallará su siguiente tiro libre?
- Prueba.** ¿Cuál es la probabilidad de adivinar incorrectamente cuando se hace una elección al azar en una pregunta de un examen de opción múltiple con respuestas posibles a, b, c, d y e, una de las cuales es correcta?
- Béisbol.** ¿Cuál es la probabilidad de que un jugador con porcentaje de bateo de 0.280 no obtendrá un hit en su siguiente turno al bat?
- Defectos.** ¿Cuál es la probabilidad de no obtener un fusible defectuoso cuando se elige al azar un fusible de una línea de ensamblado y 2% de los fusibles son defectuosos?

Probabilidades teóricas. Para los ejercicios 29 al 32 utilice el método teórico con el fin de determinar la probabilidad del resultado o evento dado. Indique cualquier suposición que necesite hacer.

- M&M.** Una bolsa contiene 10 M&M rojos, 15 M&M azules y 20 M&M amarillos. ¿Cuál es la probabilidad de sacar un M&M rojo? ¿Uno azul? ¿Uno amarillo? ¿Uno que no sea amarillo?
- Preguntas de examen.** El Colegio de Medicina de Nueva Inglaterra utiliza un examen de admisión con preguntas de opción múltiple, cada una con cinco posibles respuestas, sólo una de las cuales es correcta. Si usted elige aleatoriamente todas las respuestas, ¿qué calificación esperaba obtener (en términos de porcentaje)?
- Familia con tres hijos.** Suponga que selecciona aleatoriamente una familia con tres hijos. Si los nacimientos de hombres y mujeres son igualmente probables, ¿cuál es la probabilidad de que la familia tenga cada uno de lo siguiente?
 - Tres mujeres.
 - Dos niños y una niña.
 - Una niña, un niño y un niño, en ese orden.
 - Al menos una mujer.
 - Al menos dos hombres.

32. Familia con cuatro hijos. Suponga que selecciona al azar a una familia con cuatro hijos y que los nacimientos de hombres y mujeres son igualmente probables.

- ¿Cuántos órdenes de nacimiento son posibles? Liste todos.
- ¿Cuál es la probabilidad de que la familia tenga cuatro hombres?
- ¿Cuál es la probabilidad de que la familia tenga un niño, una niña, un niño y una niña, en ese orden?
- ¿Cuál es la probabilidad de que la familia tenga dos mujeres y dos hombres, en cualquier orden?

Probabilidades con frecuencias relativas. Utilice el método de la frecuencia relativa para estimar las probabilidades en los ejercicios 33 a 36.

33. Pronóstico de clima. Después de registrar el pronóstico del clima de 30 días, concluye que hay un pronóstico correcto 12 veces. ¿Cuál es la probabilidad de que su siguiente pronóstico sea correcto?

34. Inundación. Este año, ¿cuál es la probabilidad de una inundación que se repite cada 100 años?

35. Básquetbol. A mitad de temporada, una jugadora de básquetbol ha encestado 86% de sus tiros libres. ¿Cuál es la probabilidad de que acierte su siguiente tiro libre?

36. Cirugía. En una clínica de 73 pacientes con síndrome del túnel carpiano tratados con cirugía, 67 han tenido tratamientos exitosos (con base en información de “Splinting vs. Surgery in the Treatment of Carpal Tunnel Syndrome” de Gerritsen, *et al.*, *Journal of the American Medical Association*, volumen 288, número 10). ¿Cuál es la probabilidad de que el siguiente tratamiento con cirugía sea exitoso?

37. Probabilidades de personas de la tercera edad. En el año 2000 había 34.7 millones de personas de más de 65 años de edad en una población de 281 millones en Estados Unidos. En el año 2050 habrá, según estimaciones, 78.9 millones de personas mayores de 65 años en una población de 394 millones en Estados Unidos. ¿Sus posibilidades de encontrarse al azar con una persona de más de 65 años son mayores en 2000 o en 2050? Explique.

38. Edad en el primer matrimonio. La tabla siguiente proporciona los porcentajes de mujeres y hombres casados por primera vez en varias categorías de edad (Oficina del Censo de Estados Unidos).

	Menor de 20	20-24	25-29	30-34	35-44	45-64	Más de 65
Mujeres	16.6	40.8	27.2	10.1	4.5	0.7	0.1
Hombres	6.6	36.0	34.3	14.8	7.1	1.1	0.1

- ¿Cuál es la probabilidad de que una mujer casada, encontrada aleatoriamente, esté casada por primera vez, entre las edades de 35 y 44?

b. ¿Cuál es la probabilidad de que un hombre casado, encontrado aleatoriamente, se casó por primera vez, antes de cumplir 20 años?

c. Construya una gráfica de barras juntas que representen a hombres y a mujeres.

39. Distribución de probabilidad con cuatro monedas.

a. Construya una tabla similar a la 6.2, que muestre todos los resultados posibles del lanzamiento de cuatro monedas a la vez.

b. Construya una tabla similar a la 6.4, que muestre la distribución de probabilidad para los eventos 4 caras, 3 caras, 2 caras, 1 cara y 0 caras.

c. ¿Cuál es la probabilidad de obtener 2 caras y 2 cruces cuando lanza cuatro monedas a la vez?

d. ¿Cuál es la probabilidad de obtener cualquier cosa excepto 4 caras cuando lanza cuatro monedas al mismo tiempo?

e. ¿Cuál evento es más probable que ocurra?

40. Distribución de la lotería de Colorado. El histograma en la figura 6.8 muestra la distribución de 5964 números de la lotería de Colorado (los valores posibles van de 1 a 42).

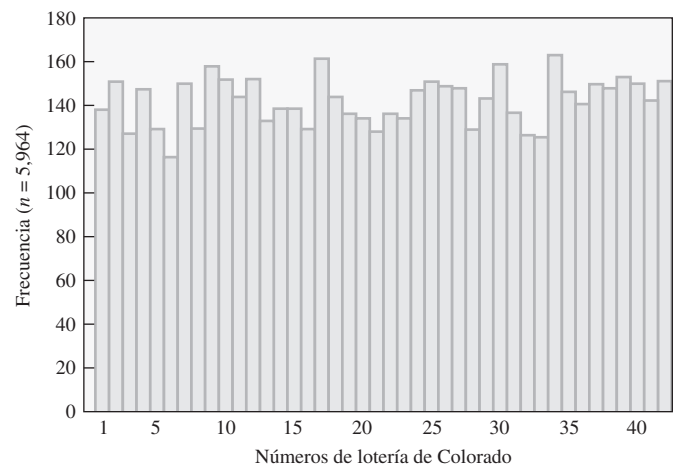


Figura 6.8 Distribución de la lotería de Colorado.

a. Si cada número de la lotería es extraído al azar, ¿cuál es la probabilidad de obtener un número cualquiera?

b. Con base en el histograma, ¿cuál es la probabilidad de frecuencia relativa del número que aparece con mayor frecuencia?

c. Con base en el histograma, ¿cuál es la probabilidad de frecuencia relativa del número que aparece con menos frecuencia?

d. Comente sobre las desviaciones de las probabilidades empíricas de las probabilidades esperadas. ¿Diría que estas desviaciones son significativas?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 6 en www.aw.com/bbt.

- 41. Grupos de sangre.** Los cuatro grupos sanguíneos principales se designan con A, B, AB y O. En cada grupo hay dos tipos Rh; positivo y negativo. Usando recursos de una biblioteca o internet, encuentre información sobre la frecuencia relativa de estos grupos sanguíneos, incluyendo los tipos Rh. Construya una tabla que muestre la probabilidad de encontrarse con alguien en cada una de las ocho combinaciones de grupos sanguíneos.
- 42. Edad y género.** Las proporciones de hombres y mujeres en la población cambian con la edad. Usando datos actuales de un sitio web, construya una tabla que muestre la probabilidad de encontrarse con un hombre o una mujer en cada una de estas categorías: 0-5, 6-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, mayor a 80.
- 43. Probabilidades de una tachuela.** Tome una tachuela común y practique lanzándola sobre una superficie plana. Observe que existen dos resultados diferentes. La tachuela apunta hacia abajo o apunta hacia arriba.
 - a. Lance la tachuela 50 veces y registre los resultados.
 - b. Con base en éstos, proporcione las probabilidades de frecuencia relativa de los dos resultados.
 - c. Si es posible, pida a otras personas que repitan el proceso. ¿Qué tan bien coinciden sus probabilidades?
- 44. Experimento con tres monedas.** Lance tres monedas a la vez, 50 ocasiones, y registre los resultados en términos del número de caras. Con base en sus observaciones, proporcione las probabilidades de frecuencia relativa de los resultados. ¿Coinciden con las probabilidades teóricas? Explique y analice sus resultados.
- 45. Aleatorización de una encuesta.** Suponga que quiere realizar una encuesta que incluya una pregunta delicada, que no todos los participantes elijan responder de manera honesta (por ejemplo, una pregunta que involucre mentir sobre los impuestos o el uso de drogas). A continuación indicamos una manera de realizar la encuesta y proteger la identidad del encuestado. Supondremos que la pregunta delicada requiere una respuesta *sí* o *no*. Primero, pida a

todos los participantes que lancen una moneda. Luego proporcione las instrucciones siguientes:

- Si obtiene cara, entonces responda la pregunta señuelo (*sí/no*). ¿Usted nació en un día par del mes?
- Si obtiene una cruz, entonces responda la pregunta real de la encuesta (*sí/no*).

Después de que todos los participantes respondieron *sí* o *no* a la pregunta que se les asignó, cuente el número total de *sí* y de *no*.

- a. Seleccione una pregunta que podría no producir respuestas totalmente honestas y realice una encuesta en su grupo usando esta técnica.
- b. Proporcione sólo el número total de respuestas *sí* y de *no* a ambas preguntas, explique cómo puede estimar el número de personas que respondieron *sí* a la pregunta real.
- c. ¿Los resultados calculados en la parte b son exactos? Explique.
- d. Suponga que la pregunta señuelo se reemplazó por estas instrucciones. Si usted obtiene cara, entonces responda *sí*. ¿Aún puede determinar el número de personas que respondieron *sí* y *no* a la pregunta real?



EN LAS NOTICIAS



- 46. Probabilidades teóricas.** Encuentre una noticia o reportaje de investigación que cite una probabilidad teórica. Redacte un comentario crítico de un párrafo.
- 47. Probabilidades de frecuencia relativa.** Encuentre una noticia o reportaje de investigación que haga uso de una probabilidad de frecuencia relativa (o probabilidad empírica). Redacte un comentario crítico de un párrafo.
- 48. Probabilidades subjetivas.** Encuentre una noticia o un reportaje de investigación que se refiera a una probabilidad subjetiva. Redacte un comentario crítico de un párrafo.
- 49. Distribuciones de probabilidad.** Encuentre una noticia o un reportaje de investigación que cite o haga uso de una distribución de probabilidad. Redacte un comentario crítico de un párrafo.

6.3 Probabilidad con grandes números

Cuando hacemos observaciones o medidas que tienen resultados aleatorios, es imposible predecir cualquier resultado individual. Sólo podemos establecer la *probabilidad* de un resultado individual. Sin embargo, cuando hacemos muchas observaciones o mediciones, esperamos que la distribución de los resultados muestre algún patrón o regularidad. En esta sección utilizamos la idea de probabilidad para formular enunciados útiles acerca de muestras grandes. En el proceso veremos una de las relaciones más importantes entre probabilidad y estadística.

Las ideas falsas predominantes en todas las clases de la comunidad respecto a casualidad y suerte ilustran la verdad de que el consentimiento común apunta casi por necesidad al error.

—Richard Proctor,
Chance and Luck
(libro de texto de 1887)

La ley de los grandes números

Si usted lanza una moneda una vez, no puede predecir exactamente cómo caerá; sólo puede establecer que la probabilidad cara es 0.5. Si lanza la moneda 100 veces, todavía no puede predecir de manera precisa cuántas caras ocurrirán. Sin embargo, de manera razonable puede esperar obtener cara *cerca de* 50% de las veces. Si lanza la moneda 1 000 veces, puede esperar que la proporción de caras sea más cercana a 50%. En general, entre más veces lance la moneda, el porcentaje será más cercano a exactamente 50%. La idea de que números grandes de eventos puedan mostrar algún patrón aunque los eventos individuales sean impredecibles se denomina **ley de los grandes números** (o *ley de los promedios*).

Ley de los grandes números

La **ley de los grandes números** (o ley de los promedios) se aplica a un proceso para el cual la probabilidad de un evento A es $P(A)$ y los resultados de ensayos repetidos no dependen de los resultados de ensayos previos (son *independientes*). Esto es: *si el proceso se repite en muchos ensayos, la proporción de los ensayos en los que el evento A ocurre será cercano a la probabilidad $P(A)$. Entre mayor sea el número de ensayos, la proporción debe ser más cercana a $P(A)$.*

Podemos ilustrar la ley de los grandes números con un experimento de tiro de un dado. La probabilidad de un 1 en un solo tiro es $P(1) = 1/6 = 0.167$. Para evitar el tedio de tirar el dado muchas veces, podemos dejar que una computadora *simule* tiros aleatorios del dado. La figura 6.9 muestra los resultados de una simulación de computadora del tiro de un solo dado 5 000 veces. El eje horizontal proporciona el número de tiros, y la altura de la curva da la proporción de unos. Aunque la curva oscila cuando el número de tiros es pequeño, para números grandes de tiros la proporción de unos se aproxima a la probabilidad 0.167, justo como lo predijo la ley de los grandes números.

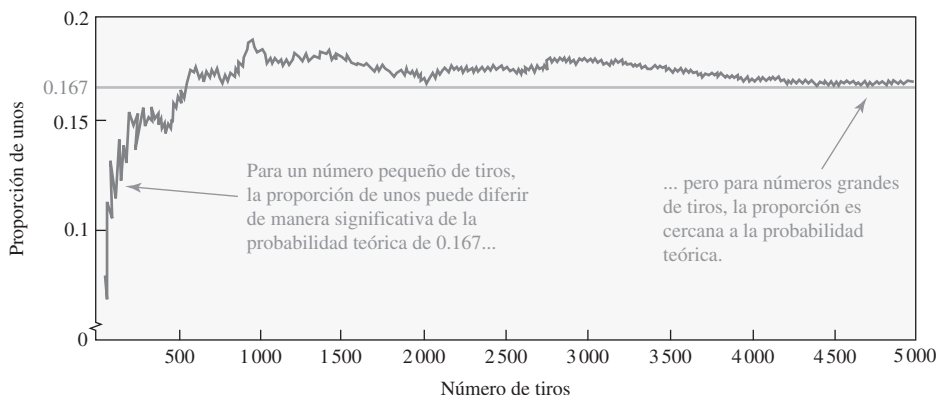


Figura 6.9 Resultados de una simulación de cómputo del tiro de un dado. Cuando el número de tiros se hace grande, la proporción de unos se hace cercana a la probabilidad teórica de 1 en un solo tiro, que es 0.167.

EJEMPLO 1 Ruleta

Una ruleta tiene 38 números: 18 números negros, 18 números rojos y los números 0 y 00 en verde. (Suponga que todos los resultados —los 38 números— tienen igual probabilidad).

- ¿Cuál es la probabilidad de obtener un número rojo en cualquier giro?
- Si los clientes en un casino giran la ruleta 100 000 veces, ¿cuántas veces esperaríamos un número rojo?

Solución

- La probabilidad teórica de obtener un número rojo en un giro es

$$P(A) = \frac{\text{número de maneras en que ocurre rojo}}{\text{número total de resultados}} = \frac{18}{38} = 0.474$$

- La ley de los grandes números nos dice que entre más y más veces se gire la ruleta, la proporción de veces que un número rojo aparece debe ser más cercana a 0.474. En 100 000 intentos, la ruleta caerá en rojo cerca de 47.4% de las veces o alrededor de 47 400 veces.

Valor esperado

Suponga que la compañía Seguro Total vende un tipo especial de seguro en el que promete pagarle \$100 000 en el evento de que deba abandonar su trabajo a consecuencia de una enfermedad grave. Con base en datos de reclamos anteriores, la probabilidad de que al asegurado se le pague por pérdida de trabajo es 1 en 500. ¿La compañía de seguros esperaría obtener una utilidad si vende las pólizas en \$250 cada una?

Si Seguro Total sólo vende unas cuantas pólizas, la utilidad o pérdida es impredecible. Por ejemplo, la venta de 100 pólizas por \$250 cada una le genera un ingreso de $100 \times \$250 = \$25\,000$. Si ninguno de los 100 tenedores de póliza presenta una demanda, la compañía hará una considerable utilidad. Por otra parte, si Seguro Total debe pagar \$100 000 incluso a un solo asegurado, se enfrentaría a una gran pérdida.

En contraste, si Seguro Total vende un número grande de pólizas, la ley de los grandes números nos dice que la proporción de pólizas con reclamos que tendrán que pagarse debe ser muy cercana a una probabilidad de 1 en 500 para una póliza individual. Por ejemplo, si la compañía vende un millón de pólizas, esperaríamos que el número de asegurados que cobrarían un reclamo de \$100 000 sería cercano a

$$\underbrace{1\,000\,000}_{\text{número de pólizas}} \times \underbrace{\frac{1}{500}}_{\text{probabilidad de reclamo de \$100 000}} = 2\,000$$

El pago de estos 2 000 reclamos costaría

$$2\,000 \times \$100\,000 = \$200 \text{ millones}$$

Este costo es un *promedio* de \$200 para cada una del millón de pólizas, lo cual significa que si las pólizas se venden en \$250 cada una, la compañía debe esperar generar un promedio de $\$250 - \$200 = \$50$ por póliza. A este promedio le llamamos el **valor esperado** para cada póliza; observe que es “esperado” sólo si la compañía vende un gran número de pólizas.

Definición

El **valor esperado** de una variable es el promedio ponderado de todos sus posibles eventos. Puesto que es un promedio, esperaríamos encontrar el “valor esperado” sólo cuando haya un número grande de eventos, de modo que entra en escena la ley de los grandes números.

A propósito...

El huracán Katrina (en 2005) fue la catástrofe natural más costosa en la historia de Estados Unidos, responsable de casi 2 000 muertes y daños estimados en más de 80 000 millones de dólares.



Podemos encontrar el mismo valor esperado con un procedimiento más formal. El ejemplo del seguro implica dos eventos distintos, cada uno con una *probabilidad* y *valor* particulares para la compañía:

1. En el evento de que una persona compre una póliza, el valor para la compañía es el precio \$250 de la póliza. La probabilidad de este evento es 1, ya que todo aquel que compra una póliza paga \$250.
2. En el evento que a una persona se le pague un reclamo, el valor para la compañía es $-\$100\,000$; es negativo ya que la compañía pierde \$100 000 en este caso. La probabilidad de este evento es $1/500$.

Ahora multiplicamos el valor de cada evento por su probabilidad y sumamos los resultados para determinar el valor esperado de cada póliza de seguro:

$$\begin{aligned} \text{valor esperado} &= \underbrace{\$250}_{\text{valor de venta de la póliza}} \times \underbrace{1}_{\text{probabilidad de obtener \$250 en la venta}} + \underbrace{(-\$100\,000)}_{\text{valor del reclamo}} \times \underbrace{\frac{1}{500}}_{\text{probabilidad de pago del reclamo}} \\ &= \$250 - \$200 = \$50 \end{aligned}$$

Esta utilidad esperada de \$50 por póliza es la misma respuesta encontrada anteriormente. Observe que la utilidad asciende a \$50 millones en ventas de 1 millón de pólizas.

Cálculo del valor esperado

Considere dos eventos, cada uno con su propio valor y probabilidad. El **valor esperado** es

$$\text{valor esperado} = \left(\text{valor del evento 1} \right) \times \left(\text{probabilidad del evento 1} \right) + \left(\text{valor del evento 2} \right) \times \left(\text{probabilidad del evento 2} \right)$$

Esta fórmula puede ampliarse a cualquier número de eventos mediante la inclusión de más términos en la suma.

NOTA TÉCNICA

Existen formas alternas para calcular el valor esperado. En el ejemplo del seguro podríamos definir el evento 1 como una póliza sin reclamo y el evento 2 como una póliza a la cual la compañía paga un reclamo de \$100 000. El evento 1 tiene un valor para la compañía de \$250 (el precio de venta de una póliza) y una probabilidad $499/500$. El evento 2 tiene un valor para la compañía de \$250 (el precio de la póliza) *menos* los \$100 000 que paga por un reclamo, o $-\$99\,750$, y una probabilidad de $1/500$. Observe que este método alternativo aún proporciona el mismo valor esperado de \$50:

$$\begin{aligned} &\left(\$250 \times \frac{499}{500} \right) \\ &+ \left(-\$99\,750 \times \frac{1}{500} \right) = \$50 \end{aligned}$$

Algunos estadísticos prefieren este método alternativo porque la suma de las probabilidades de los eventos es 1.

UN MOMENTO DE REFLEXIÓN

¿La compañía de seguros debe esperar obtener una utilidad de \$50 en cada póliza individual? ¿Debe esperar una utilidad de \$50 000 en 1 000 pólizas? Explique.

EJEMPLO 2 Esperanzas en loterías

Suponga que un billete de lotería de \$1 tiene las probabilidades siguientes: 1 en 5 para ganar un boleto gratis (valor de \$1), 1 en 100 para ganar \$5, 1 en 100 000 para ganar \$1 000 y 1 en 10 millones para ganar \$1 millón. ¿Cuál es el valor esperado de un billete de lotería? Analice las implicaciones. (Nota: los ganadores *no* obtienen el \$1 que gastaron en el billete).

Solución La manera más sencilla de proceder es construir una tabla (a continuación) de todos los eventos relevantes con sus valores y probabilidades. Nosotros calculamos el valor esperado de un billete de lotería para *usted*; así, el precio del billete tiene un valor negativo ya que le cuesta su dinero, mientras que los valores de las ganancias son positivos.

Evento	Valor	Probabilidad	Valor × probabilidad
Compra de billete	−\$1	1	$(-\$1) \times 1 = -\1.00
Gana billete gratis	\$1	$\frac{1}{5}$	$\$1 \times \frac{1}{5} = \0.20
Gana \$5	\$5	$\frac{1}{100}$	$\$5 \times \frac{1}{100} = \0.05
Gana \$1 000	\$1 000	$\frac{1}{100\,000}$	$\$1\,000 \times \frac{1}{100\,000} = \0.01
Gana \$1 millón	\$1 000 000	$\frac{1}{10\,000\,000}$	$\$1\,000\,000 \times \frac{1}{10\,000\,000} = \0.10
			Suma de la última columna: −\$0.64

El valor esperado es la suma de todos los producto *valor* $\times$ *probabilidad*, que al final de la tabla muestra que es $-\$0.64$. Así, promediando sobre todos los billetes, debe esperar perder 64¢ por cada billete de lotería que compre. Si compra, digamos, 1 000 billetes, debe esperar perder $1\,000 \times \$0.64 = \640 .

UN MOMENTO DE REFLEXIÓN

Muchos gobiernos estatales utilizan loterías para financiar causas justas tal como parques, recreación y educación. Las loterías también tienden a mantener los impuestos estatales en niveles bajos. Por otra parte, la investigación muestra que las loterías son pagadas por personas con bajos ingresos. ¿Usted considera que las loterías son una buena política social? ¿Cree que las loterías son una política económica buena?

DILBERT® by Scott Adams



© 1994 Scott Adams, Inc. Distribuido por United Feature Syndicate. Reimpreso con permiso. Todos los derechos reservados.

La falacia del jugador

Considere un juego sencillo que implica el lanzamiento de una moneda. Usted gana \$1 si la moneda cae cara y pierde \$1 si cae cruz. Suponga que lanza la moneda 100 veces y obtiene 45 caras y 55 cruces, lo que le crea una deuda de \$10. ¿"Debe" tener una racha de mejor suerte?

Quizá reconozca que la respuesta es *no*: su mala suerte pasada no tiene relación con sus oportunidades futuras. Sin embargo, muchos jugadores —en especial los jugadores compulsivos— suponen justamente lo opuesto. Ellos creen que cuando su suerte ha sido mala, debe haber un cambio. Esta creencia errónea con frecuencia se denomina la **falacia del jugador** (o la *ruina del jugador*).

Definición

La **falacia del jugador** es la falsa creencia de que una racha de mala suerte le "garantiza" a una persona una racha de buena suerte.

Alguien que apuesta cualquier parte de su fortuna, por pequeña que sea, en un juego matemáticamente no justo, actúa de manera irracional. . .

La imprudencia de un jugador será lo más grande de su fortuna que expone a un juego de azar.

—Daniel Bernoulli,
matemático del siglo XVIII

Una razón por la que la gente sucumbe a la falacia del jugador es por una mala comprensión de la ley de los grandes números. En el juego del lanzamiento de una moneda, la ley de los grandes números nos dice que la proporción de caras tiende a ser más cercana a 50% para un número grande de lanzamientos. Pero esto *no* significa que usted posiblemente se recupere de pérdidas anteriores. Para ver por qué, estudie la tabla 6.7, la cual muestra los resultados de simulaciones de computadora de un número grande de lanzamientos de una moneda. Observe que cuando el número de lanzamientos aumenta, el porcentaje de caras se hace más cercano a exactamente 50%, justo como predice la ley de los grandes números. Sin embargo, la última columna muestra que la *diferencia* entre el número de caras y el número de cruces continúa creciendo, esto significa que en este conjunto particular de simulaciones, las pérdidas (la diferencia entre el número de caras y el número de cruces) crecen aun cuando la proporción de caras se aproxime a 50%.

Tabla 6.7 Resultados de ensayos de lanzamiento de una moneda			
Número de lanzamientos	Número de caras	Porcentaje de caras	Diferencia entre los números de caras y cruces
100	45	45%	10
1000	470	47%	60
10 000	4 950	49.5%	100
100 000	49 900	49.9%	200

La figura 6.10 muestra este resultado de manera gráfica para una simulación de computadora de una moneda lanzada 1 000 veces. Aunque el porcentaje de caras tiende a 50% a lo largo de los 1 000 lanzamientos, vemos desviaciones de la igualdad en el número de caras y cruces.

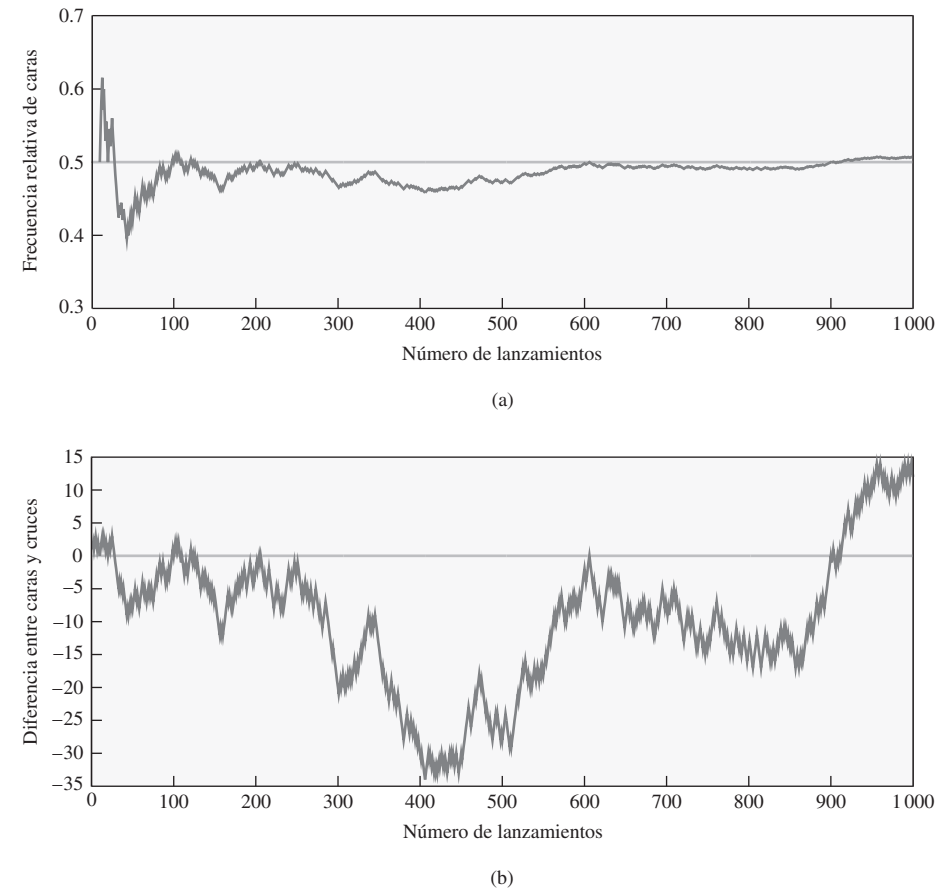


Figura 6.10 Una simulación de computadora de 1000 lanzamientos de una moneda. (a) Esta gráfica muestra cómo la proporción de caras cambia cuando el número de lanzamientos aumenta. Observe que la proporción se aproxima a 0.5, como esperamos de la ley de los grandes números. (b) Esta gráfica muestra cómo la *diferencia* entre los números de caras y cruces cambia cuando el número de lanzamientos aumenta. Observe que la diferencia, en realidad, puede crecer y hacerse grande con más lanzamientos, a pesar de que la proporción de caras se aproxime a 0.5.

EJEMPLO 3 Pérdidas continuas

Usted juega al lanzamiento de moneda y gana \$1 por cada cara y pierde \$1 por cada cruz. Después de 100 lanzamientos, usted tiene \$10 en pérdidas porque tiene 45 caras y 55 cruces. Usted continúa jugando hasta que ha lanzado la moneda 1 000 veces, en las cuales ha obtenido 480 caras y 520 cruces. ¿Este resultado es consistente con lo que esperamos de la ley de los grandes números? ¿Ha recuperado algo de sus pérdidas? Explique.

Solución La proporción de caras en sus primeros 100 lanzamientos fue 45%. Después de 1 000 lanzamientos la proporción de caras ha aumentado a 480 de 1 000 o 48%. Puesto que la proporción de caras se ha movido más cerca de 50%, el resultado es consistente con lo que esperamos de la ley de los grandes números. Sin embargo, ahora ganó \$480 (por las 480 caras) y perdió \$520 (de las 520 cruces), para una pérdida neta de \$40. Por tanto, sus pérdidas han *aumentado*, a pesar de que la proporción de caras aumentó y es más cercana a 50%.

Rachas

Otro mal entendido que contribuye a la falacia del jugador involucra las esperanzas de rachas. Suponga que lanza una moneda seis veces y ve el resultado HHHHHH (todas caras). Entonces lanza la moneda seis veces más y ve el resultado HTTHTH. La mayoría de la gente diría que el último resultado es “natural” mientras que la racha de todas caras es sorprendente. Pero, de hecho, ambos resultados son igualmente probables. El número total de resultados para seis monedas es $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 64$, y cada resultado individual tiene la misma probabilidad $1/64$.

Además, suponga que acaba de obtener seis caras y tiene que apostar el resultado del siguiente lanzamiento. Podría pensar que, dada la racha de caras, una cruz “debe” salir en el lanzamiento siguiente. Pero la probabilidad de una cara o una cruz en el siguiente lanzamiento aún es 0.50; la moneda no tiene memoria de los lanzamientos previos.

UN MOMENTO DE REFLEXIÓN

¿Es más o menos probable que una familia de seis hijos, todos varones, tenga un varón en el siguiente nacimiento? ¿Una jugadora de básquetbol, que ha anotado 20 tiros libres consecutivos, es más o menos probable que anote su siguiente tiro libre? ¿El clima de un día es independiente del clima en el siguiente (como se supone en el ejemplo)? Explique.

EJEMPLO 4 Planeación para la lluvia

Un granjero sabe que en esta época del año en su región del país, la probabilidad de lluvia en un día dado es 0.5. No ha llovido durante 10 días, y él necesita decidir si empieza a regar. ¿Tiene justificación para posponer el riego porque él debe tener un día lluvioso?

Solución Los 10 días de sequía no son esperados y, como un jugador, el granjero está teniendo una “racha de mala suerte”. Sin embargo, si suponemos que los eventos de clima son independientes de un día al siguiente, entonces es una falacia esperar que la probabilidad de lluvia sea mayor o menor que 0.5.

A propósito...

La misma clase de razonamiento acerca de rachas se aplica a los números seleccionados de la lotería. No hay combinaciones especiales de los números de la lotería que sean más probables de salir que otras combinaciones. Sin embargo, la gente compra combinaciones especiales (tales como 1, 2, 3, 4, 5, 6) con más frecuencia que combinaciones ordinarias. Por tanto, si gana una lotería con una combinación ordinaria, es menos probable que divida el premio con otros, con el efecto neto de que su premio será mayor.

Sección 6.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Ley de los grandes números.** Con sus palabras describa la ley de los grandes números.
- 2. Valor esperado.** Un genetista calcula el número esperado de niñas en 5 nacimientos y obtiene el resultado de 2.5 niñas. Él redondea el resultado a 3 niñas, bajo el razonamiento que es imposible tener 2.5 niñas en 5 nacimientos. ¿Es correcto ese razonamiento? ¿Por qué sí o por qué no?
- 3. Estrategia de juego.** Un jugador profesional de blackjack en el casino Veneciano ha perdido en cada una de sus pri-

meras 10 apuestas. Él empieza a poner apuestas más grandes, razonando que su proporción actual de ganados (que es 0) aumentará para acercarse al número promedio de ganados. ¿Su estrategia de apuesta es razonable? ¿Es correcto su razonamiento? Explique.

- 4. Falacia del jugador.** Con sus palabras describa la falacia del jugador.

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Estrategia en la ruleta.** Steve acaba de perder \$200 colocando 40 apuestas consecutivas al número 13 en una ruleta y no ha ganado en ninguna de ellas. Él razona que, de acuerdo con la ley de los grandes números, su racha de perdidos debe terminar por lo que su proporción de ganados debe ser cercana a $1/38$. Por tanto, él decide cubrir sus pérdidas colocando mayores apuestas a los siguientes giros de la ruleta.
- 6. Lotería.** Kelly trabaja en una tienda de conveniencia y tiene acceso a una máquina de lotería. Ella razona que si compra un billete para cada posible combinación de números en la lotería estatal, definitivamente ganará el premio mayor, por lo que planea reunir suficiente dinero para pagar todos esos billetes.
- 7. Lotería.** Kim compra un billete de lotería estatal y selecciona los números 1, 2, 3, 4, 5 y 6. Ella razona que esta combinación tiene la misma oportunidad de ganar que cualquier otra combinación.
- 8. Lanzamiento de moneda.** Cuando una moneda justa se lanza 10 veces, el resultado de 10 caras nunca ocurrirá.

Conceptos y aplicaciones

- 9. Comprensión de la ley de los grandes números.** Suponga que lanza una moneda 10 000 veces. ¿Debe esperar obtener exactamente 5 000 caras? ¿Por qué sí o por qué no? ¿Qué dice la ley de los grandes números acerca de los resultados que quizás obtendrá?
- 10. Conductor rápido.** Una persona que tiene el hábito de conducir rápido nunca ha sufrido un accidente o una multa de tránsito. ¿Qué significa decir que “la ley de los promedios lo atrapará”? ¿Es cierto? Explique.
- 11. ¿Debe jugar?** Suponga que alguien le da posibilidades de 5 a 1 que no puede obtener dos números pares con el tiro de dos dados. Esto significa que usted gana \$5 si tiene éxito y pierde \$1 si falla. ¿Cuál es el valor esperado de este juego para usted? ¿Debe esperar ganar o perder el valor esperado en el primer juego? ¿Qué puede esperar si juega 100 veces? Explique. (La tabla 6.5 en la página 246 le ayudará a encontrar las probabilidades requeridas).
- 12. Lotería Elija 4 de Kentucky.** Si usted apuesta \$1 en la lotería Elija 4 de Kentucky, o bien pierde \$1 o bien gana \$4 999. (El premio al ganador es \$5 000, pero su apuesta de \$1 no se le regresa, así que la ganancia neta es \$4 999). El juego trata de la selección de un número de cuatro dígitos entre 0000 y 9999. ¿Cuál es la probabilidad de ganar? Si usted apuesta \$1 en 1234, ¿cuál es el valor esperado de su ganancia o pérdida?
- 13. Puntos extra en fútbol americano.** Los equipos de fútbol americano tienen la opción de intentar anotar 1 o 2 puntos extra después de un touchdown (anotación de seis puntos). Obtienen un punto si patean el balón y pasa entre los postes o 2 puntos corriendo o pasando el balón y llegando a la zona de anotación. Para un año reciente en la NFL (Liga Nacional de Fútbol, por sus siglas en inglés), patadas de 1 punto fueron exitosas 94% de las veces, mientras que intentos de 2 puntos fueron exitosos sólo 37% de las veces. En cualquier caso, la falla significa 0 puntos. Calcule los valores esperados de intentos de 1 y de 2 puntos. Con base en estos valores esperados, ¿cuál de las dos opciones tiene más sentido en la mayoría de los casos? ¿Considera que existen circunstancias en las que se debe tomar una decisión diferente de la que sugieren los valores esperados? Explique.
- 14. Reclamos en seguros.** Un actuario en una compañía de seguros estima, a partir de datos existentes, que en una póliza de \$1 000, un promedio de 1 en 100 asegurados presentarán un reclamo por \$20 000, un promedio de 1 en 200 asegurados presentarán un reclamo por \$50 000 y un promedio de 1 en 500 asegurados presentarán un reclamo por \$100 000.
- ¿Cuál es el valor esperado para la compañía por cada póliza vendida?
 - Si la compañía vende 100 000 pólizas, ¿puede esperar una utilidad? Explique las suposiciones de este cálculo.
- 15. Tiempo esperado de espera.** Suponga que usted llega aleatoriamente a una parada de autobús, así que todos los tiempos de llegada son igualmente probables. El autobús llega regularmente cada 30 minutos sin retraso (digamos, a la hora y a la media hora) ¿Cuál es el valor esperado de su tiempo de espera? Explique cómo obtuvo su respuesta.
- 16. Lotería Powerball.** La lotería multiestatal Powerball anuncia los premios siguientes y las probabilidades de ganar para un billete sencillo de \$1. Suponga que, en una semana, el gran premio (o premio gordo) tiene un valor de \$30 millones. Observe que hay más de una forma de obtener algunos de los premios monetarios (por ejemplo, hay dos formas de ganar \$100), por lo que la tabla proporciona la probabilidad para cada caso. ¿Cuál es el valor esperado de las ganancias para un billete sencillo de lotería? Si usted gasta \$365 por año en la lotería, ¿cuánto puede esperar ganar o perder?

Premio	Probabilidad
Premio gordo	1 en 80 089 128
\$100 000	1 en 1 953 393
\$5 000	1 en 364 042
\$100	1 en 8 879
\$100	1 en 8 466
\$7	1 en 207
\$7	1 en 605
\$4	1 en 188
\$3	1 en 74

- 17. Lotería Big Game.** La lotería multiestatal Big Game anuncia los premios y probabilidades siguientes de ganar para un billete sencillo de \$1. El premio mayor es variable, pero suponga que tiene un valor promedio de \$3 millones.

Observe que el mismo premio puede darse por dos resultados con diferentes probabilidades. ¿Cuál es el valor esperado de un billete sencillo de lotería? Si usted gasta \$365 al año en la lotería, ¿cuánto puede esperar ganar o perder?

Premio	Probabilidad
Premio gordo	1 en 76 275 360
\$150 000	1 en 2179 296
\$5 000	1 en 339 002
\$150	1 en 9 686
\$100	1 en 7 705
\$5	1 en 220
\$5	1 en 538
\$2	1 en 102
\$1	1 en 62

18. Valor esperado en la ruleta. Cuando usted hace una apuesta de \$5 al número 7 en la ruleta en el casino Veneziano en las Vegas, tiene una probabilidad $37/38$ de perder \$5 y tiene una probabilidad $1/38$ de tener una ganancia neta de \$175. (El premio es \$180, pero no se le regresa su apuesta de \$5, por lo que la ganancia neta es \$175). Si apuesta \$5 a que el resultado es un número impar, la probabilidad de perder \$5 es $20/38$ y la probabilidad de obtener una ganancia neta de \$5 es $18/38$. (Si usted apuesta \$5 a un número impar y gana, se le dan \$10 que incluyen su apuesta, por lo que la ganancia neta es de \$5).

- Si apuesta \$5 al número 7, ¿cuál es su valor esperado?
- Si apuesta \$5 a que el resultado es un número impar, ¿cuál es su valor esperado?
- ¿Cuál de estas opciones es mejor: apostar al 7, apostar a impar o no apostar? ¿Por qué?

19. Valor esperado en dados en un casino. Cuando da a un casino \$5 para una apuesta en la “línea de pase” en un juego de dados, hay una probabilidad $251/495$ de que pierda \$5 y una probabilidad $244/495$ de obtener una ganancia neta de \$5. (Si usted gana, el casino le da sus \$5 y usted conserva su apuesta de \$5, por lo que su ganancia neta es de \$5). ¿Cuál es su valor esperado? A largo plazo, ¿cuánto pierde por cada dólar que apuesta?

20. Tamaño medio de un hogar. Se estimó que 57% de los estadounidenses viven en hogares con 1 o 2 personas, 32% viven en hogares con 3 o 4 personas y 11% viven en hogares con 5 o más personas. Explique cómo encontraría el número esperado de personas en un hogar estadounidense. ¿Cómo está relacionado esto con el tamaño medio de un hogar?

21. Psicología de los valores esperados. En 1953 un economista francés, Maurice Allais, realizó una encuesta de cómo valora el riesgo la gente. A continuación están dos escenarios que él utilizó, cada uno pide a la gente que elija entre dos opciones.

Decisión 1:

Opción A: 100% de oportunidad de ganar \$1 000 000.

Opción B: 10% de oportunidad de ganar \$2 500 000; 89% oportunidades de ganar \$1 000 000 y 1% de no ganar.

Decisión 2:

Opción A: 11% de posibilidades de ganar \$1 000 000 y 89% de ganar \$0.

Opción B: 10% de posibilidades de ganar \$2 500 000 y 90% de posibilidades de ganar \$0.

Allais descubrió que para la decisión 1, la mayoría de la gente selecciona la opción A, mientras que para la decisión 2, la mayoría de la gente elige la opción B.

- Para cada decisión, determine el valor esperado de cada opción.
- ¿Las respuestas dadas en la encuesta son consistentes con los valores esperados?
- Proporcione una explicación para las respuestas en la encuesta de Allais.

22. Valor esperado para un sorteo. La revista *Reader's Digest* efectuó un sorteo en el que los premios fueron listados junto con sus posibilidades de ganar; \$1 000 000 (1 oportunidad en 90 000 000), \$100 000 (1 oportunidad en 110 000 000), \$25 000 (1 oportunidad en 110 000 000), \$5 000 (1 oportunidad en 36 667 000) y \$2 500 (1 oportunidad en 27 500 000).

- Suponiendo que no hay costo para entrar al sorteo, determine el valor esperado del monto ganado por una entrada.
- Determine el valor esperado si el costo de entrar a este sorteo es el costo de una estampilla postal. ¿Vale la pena entrar a este concurso?

23. Falacia del jugador y dados. Suponga que tira un dado con un amigo, con las reglas siguientes: por cada número par que usted obtenga, gana \$1 a su amigo; por cada número impar que obtenga, usted paga \$1 a su amigo.

- ¿Cuáles son las posibilidades de que obtenga un número par en un tiro de un dado justo? ¿Y la de un número impar?
- Suponga que en los primeros 100 tiros obtiene 45 números pares. ¿Cuánto dinero ha ganado o perdido?
- Suponga que en los segundos 100 tiros, su suerte mejora y obtiene 47 números pares. ¿Cuánto dinero ganó o perdió en los 200 tiros?
- Suponga que en los siguientes 300 tiros, nuevamente su suerte mejoró y obtiene 148 números pares. ¿Cuánto dinero ha ganado o perdido en los 500 tiros?

- e. ¿Cuáles fueron los porcentajes de números pares después de 100, 200 y 500 tiros? Explique por qué este juego ilustra la falacia del jugador.
- f. ¿Cuántos números pares habría tenido que obtener en los siguientes 100 tiros para igualar el número de pares e impares?

24. A la zaga en lanzamiento de una moneda: ¿puede alcanzar? Suponga que lanza una moneda 100 veces, obteniendo 38 caras y 62 cruces, que son 24 cruces más que caras.

- a. Explique por qué, en su siguiente tiro, la *diferencia* en los números de caras y cruces es igualmente probable que crezca a 25 o que disminuya a 23.
- b. Amplíe su explicación del inciso a para explicar por qué, si lanza la moneda 1 000 veces más, la diferencia final en los números de caras y cruces es igual de probable que sea mayor a 24 a que sea menor que 24.
- c. Suponga que continúa lanzando la moneda. Explique por qué el enunciado siguiente es verdadero: si usted se detiene en cualquier instante aleatorio, siempre será más probable tener, en total, menos caras que cruces.
- d. Suponga que usted está apostando a cara en cada lanzamiento de la moneda. Después de los primeros 100 lanzamientos, usted está en el lado perdedor (habiendo perdido 62 veces, mientras que ganó sólo 38 veces). Explique por qué, si continúa apostando, será más probable que continúe en el lado perdedor. ¿Cómo está relacionada esta respuesta con la falacia del jugador?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 6 en www.aw.com/bbt.

25. Análisis de loterías en la web. Vaya al sitio web de todas las loterías de Estados Unidos y estudie un resumen de

las posibilidades y premios en las loterías estatales y multiestatales. Seleccione cinco loterías y determine el valor esperado para las ganancias en cada caso. Analice sus resultados.

26. Ley de los grandes números. Utilice una moneda para simular 100 nacimientos: lance una moneda 100 veces, registrando los resultados, y luego convierta los resultados a géneros de bebés (cruz = niño y cara = niña). Utilice los resultados para llenar la tabla siguiente. ¿Qué sucede a la proporción de niñas en la muestra cuando el tamaño de la muestra aumenta? ¿Cómo ilustra esto la ley de los grandes números?

Número de nacimientos	10	20	30	40	50	60	70	80	90	100
Proporción de niñas										



EN LAS NOTICIAS



- 27. Ley personal de los grandes números.** Describa una situación en la que usted haya hecho uso de la ley de los grandes números, ya sea correcta o incorrectamente. ¿Por qué utilizó la ley de los grandes números en esta situación? ¿Fue útil?
- 28. Falacia del jugador en la vida.** Describa una situación en la que usted o alguien que conozca haya caído víctima de la falacia del jugador. ¿Cómo habría sido tratada la situación correctamente?
- 29. Ruina del jugador.** Describa una situación que sepa en la que alguien haya perdido casi todo por el juego. ¿Su estrategia parecía racional o ser el resultado de una adicción destructiva?

6.4 Ideas de riesgo y esperanza de vida

Una plática sencilla pero honesta con un vendedor, que le ofrece un producto nuevo:

No puedo darle detalles aún, pero ¡amará este producto! Mejorará su vida de muchas más maneras de las que pueda contar. Su único inconveniente es que eventualmente mata a todo aquel que lo use. ¿Compraría uno?

Probablemente no. Después de todo, ¿algún producto puede ser tan grandioso que moriría por él? Unas cuantas semanas después, el vendedor aparece nuevamente:

Nadie ha comprado; así que hemos realizado algunas mejoras. Su probabilidad de ser muerto por el producto ahora es de sólo 1 en 10. ¿Listo para comprar?

A pesar de la mejora, la mayoría de la gente aún mandaría a casa al vendedor, y esperaría por su inevitable regreso:

Está bien, esta vez realmente lo hemos perfeccionado. Lo hemos hecho tan seguro que tardaría 20 años en matar a tanta gente como la que vive en San Francisco. Y puede ser suyo por sólo \$20 000.

Puede sorprenderle darse cuenta que, si es como la mayoría de los estadounidenses, saltaría ante esta oferta. Después de todo, el producto es el automóvil. En verdad mejora nuestras vidas de muchas maneras y \$20 000 es un precio común. Y, dado que aproximadamente 40 000 estadounidenses mueren cada año en accidentes automovilísticos, el automóvil mata al equivalente a la población de San Francisco (aproximadamente 800 000) en alrededor de 20 años.

Como este ejemplo muestra, con frecuencia hacemos un balance entre beneficio y riesgos. En esta sección veremos cómo las ideas de probabilidad pueden ayudarnos a cuantificar el riesgo, y así permitirnos tomar decisiones informadas acerca de tales balances.

Apropósito...

Aproximadamente 115 personas mueren cada día en accidentes automovilísticos, es decir, 1 cada 13 minutos. Estos accidentes son la causa principal de muerte entre personas de edades entre 6 a 27 años. Los ocupantes de vehículos constituyen 85% de las víctimas mortales, mientras que los peatones y los ciclistas 15%.



Riesgo y viaje

¿Está usted más seguro en un automóvil pequeño o en un vehículo deportivo? ¿Los automóviles de ahora son más seguros que los de hace 30 años? Si usted necesita atravesar todo el país, ¿está más seguro volando o manejando? Para responder éstas y muchas preguntas similares, debemos cuantificar el riesgo involucrado al viajar. Luego podemos tomar decisiones apropiadas para nuestras circunstancias personales.

El riesgo de viajar con frecuencia se expresa en términos de una **tasa de accidentes** o **tasa de mortalidad**. Por ejemplo, suponga que una tasa de accidentes anual es 750 accidentes por cada 100 000 personas. Esto significa que, en un grupo de 100 000 personas, en promedio 750 tendrán un accidente en un periodo de un año. El enunciado en esencia es un valor esperado, también representa una probabilidad. Nos dice que la probabilidad de que una persona esté involucrada en un accidente (durante un año) es 750 en 100 000 o 0.0075.

Este concepto de riesgo de viaje es directo, pero aún debemos interpretar los números con cuidado. Por ejemplo, los riesgos de viaje en ocasiones se indican *por cada 100 000 personas*, como antes, pero en otras ocasiones se indican *por viaje* o *por milla*. Si utilizamos tasas de mortalidad *por viaje* para comparar los riesgos de volar y manejar, no tomamos en cuenta el hecho de que los viajes por avión por lo regular son mucho más largos que los viajes en automóvil. De manera similar, si utilizamos tasas de accidentes *por persona*, no tomamos en cuenta que la mayoría de los accidentes automovilísticos incluyen sólo lesiones menores.

EJEMPLO 1 ¿Conducir es más seguro?

La figura 6.11 muestra el número de accidentes mortales y el número total de millas manejadas (entre todos los estadounidenses) en cada año durante un periodo de más de tres décadas. En términos de tasa de mortalidad por milla conducida, ¿cómo ha cambiado el riesgo de manejo?

Solución La figura 6.11a muestra que el número anual de fatalidades disminuyó de alrededor de 52 000 en 1970 a casi 43 000 en 2004. Mientras tanto, la figura 6.11b muestra que el número de millas conducidas aumentó de casi 1 billón (1×10^{12}) a cerca de 2.9 billones (2.9×10^{12}). Por tanto, las tasas de mortalidad por milla para el inicio y el final del periodo fueron

$$1970: \frac{52\,000 \text{ muertes}}{1 \times 10^{12} \text{ millas}} \approx 5.2 \times 10^{-8} \text{ muertes por milla}$$

$$2004: \frac{43\,000 \text{ muertes}}{2.9 \times 10^{12} \text{ millas}} \approx 1.5 \times 10^{-8} \text{ muertes por milla}$$

Observe que $10^8 = 100$ millones, 5.2×10^{-8} muertes por milla es equivalente a 5.2 muertes por 100 millones de millas. Por tanto, en 34 años, la tasa de mortalidad por cada 100 millones de millas cayó de 5.2 a 1.5. Con base en esta medida, manejar se volvió mucho más seguro durante el periodo. La mayoría de los investigadores creen que las mejoras son resultado de perfeccionamientos en el diseño de los automóviles y de características de seguridad, como cinturones de seguridad y bolsas de aire, que son mucho más comunes en la actualidad.

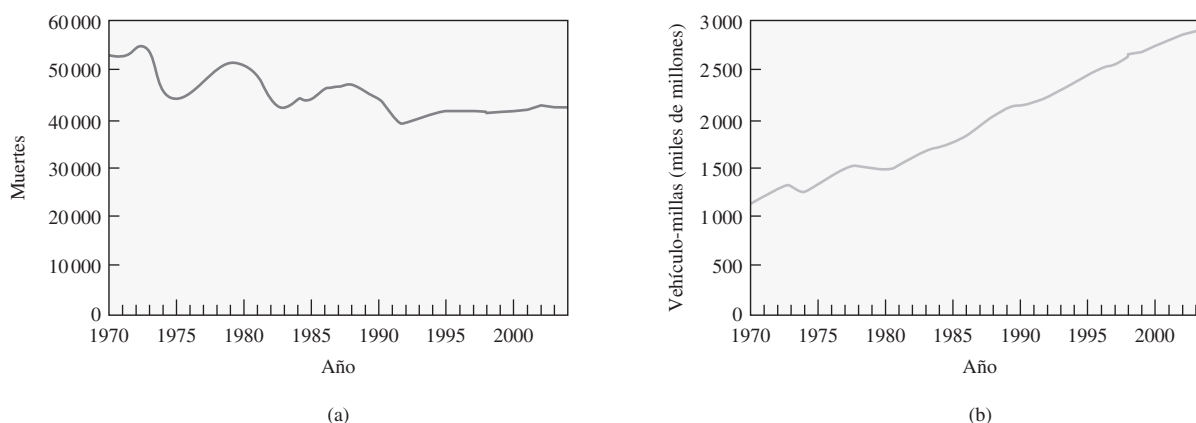


Figura 6.11 (a) Accidentes mortales de automóviles por año. (b) Millas totales manejadas durante un año. Ambos conjuntos de datos son sólo de Estados Unidos. *Fuente:* Oficina Nacional de Seguridad del Transporte.

EJEMPLO 2 ¿Qué es más seguro: volar o manejar?

Durante los 20 años anteriores, en Estados Unidos, el número promedio (media) de muertes en accidentes de aviones comerciales ha sido aproximadamente de 100 por año. (El número real varía de manera significativa de año a año). En la actualidad los pasajeros de aeroplanos en Estados Unidos viajaron cerca de 8 mil millones de millas por año. Utilice estos números para calcular la tasa de mortalidad por milla de viaje aéreo. Compare el riesgo de volar con el riesgo de conducir.

Solución Suponiendo 100 muertes y 8 mil millones de millas en un año promedio, el riesgo de viaje aéreo es

$$\frac{100 \text{ muertes}}{8 \times 10^9 \text{ millas}} \approx 1.3 \times 10^{-8} \text{ muertes por milla}$$

El costo de vivir está subiendo mientras que la posibilidad de vivir está bajando.

—Flip Wilson,
comediante

Este riesgo es equivalente a 1.3 muertes por 100 millones de millas, o ligeramente menor que el riesgo de 1.5 muertes por 100 millones de millas conducidas (vea el ejemplo 1). Observe que, puesto que el promedio de viaje aéreo cubre una distancia considerablemente mayor que el promedio de un viaje manejando, el riesgo *por viaje* es mucho más alto para viajes aéreos, aunque el riesgo *por milla* es menor.

UN MOMENTO DE REFLEXIÓN

Suponga que necesita realizar el viaje de 800 millas de Atlanta a Houston. ¿Cree más seguro volar o conducir? ¿Por qué?

Sólo aquellos que se arriesgan a ir más lejos posiblemente encuentren cómo uno puede ir más lejos.

—T. S. Eliot

Estadísticas vitales

Los datos relacionados con nacimientos y muertes de ciudadanos, con frecuencia llamadas *estadísticas vitales*, son muy importantes para comprender el balance riesgo-beneficio. Por ejemplo, las compañías de seguros utilizan estadísticas vitales para valorar riesgos y establecer tarifas. Los profesionales de la salud estudian las estadísticas vitales para evaluar el progreso médico y decidir dónde deben concentrarse los recursos de la investigación. Los demógrafos utilizan tasas de natalidad y mortalidad para pronosticar tendencias futuras de la población.

Un conjunto importante de estadísticas vitales, mostradas en la tabla 6.8, tiene que ver con causas de muerte. Estos datos son extremadamente generales; una tabla más completa clasificaría datos por edad, sexo y raza. Las estadísticas vitales con frecuencia se expresan en términos de muertes por personas o por cada 100 000 personas, que hacen más sencillo comparar tasas para diferentes años y para diferentes estados o países.

Tabla 6.8 Principales causas de muerte en Estados Unidos (en un solo año reciente)			
Causa	Muertes	Causa	Muertes
Cardiopatía	684 462	Diabetes	73 249
Cáncer	554 643	Neumonía/Influenza	65 681
Apoplejía	157 803	Alzheimer	63 343
Enfermedad pulmonar	126 128	Enfermedad renal	42 536
Accidentes	105 695	Septicemia (envenenamiento de la sangre)	34 243

Fuente: Centros de Control de Enfermedades.

Apropósito...

Entre los estudiantes a nivel universitario, el consumo de alcohol presenta uno de los más serios riesgos para la salud. El Instituto Nacional de Salud estima que el alcohol contribuye a 1 400 muertes y 500 000 lesiones entre estudiantes universitarios cada año, así como 70 000 casos de ataques sexuales.

EJEMPLO 3 Interpretación de estadísticas vitales

Suponiendo una población de Estados Unidos de 300 millones, determine y compare los riesgos por persona y por cada 100 000 personas por neumonía (e influenza) y cáncer.

Solución Encontramos el riesgo por persona dividiendo el número de muertes entre la población total de 281 millones:

Neumonía/influenza:

$$\frac{65\,681 \text{ muertes}}{300\,000\,000 \text{ personas}} \approx 0.00022 \text{ muertes por persona}$$

Cáncer:

$$\frac{554\,643 \text{ muertes}}{300\,000\,000 \text{ personas}} \approx 0.0018 \text{ muertes por persona}$$

Podemos interpretar estos números como probabilidades: la probabilidad de muerte por neumonía o influenza es alrededor de 2.2 en 10 000, mientras que la probabilidad de muerte por cáncer es alrededor de 18 en 10 000. Para ponerlas en términos de muertes por cada 100 000 personas, multiplicamos las tasas por persona por 100 000. Obtenemos una tasa de muerte por neumonía/influenza de 22 muertes por 100 000 personas, y una tasa de muerte por cáncer de 180 muertes por 100 000 personas. La probabilidad de muerte por cáncer es más de ocho veces la de muerte por neumonía o influenza.

UN MOMENTO DE REFLEXIÓN

La tabla 6.8 sugiere que la probabilidad de muerte por apoplejía es alrededor de 50% mayor que la probabilidad de muerte por accidente, pero estos datos incluyen datos de todos los grupos de edad. ¿Cómo cree que los riesgos de apoplejía y accidentes diferirían entre personas jóvenes y personas mayores? Explique.

Esperanza de vida

Ahora pasamos a la idea de *esperanza de vida*, que con frecuencia se utiliza para comparar salud global en diferentes tiempos o en diferentes países. La idea será más clara si iniciamos observando las tasas de mortalidad. La figura 6.12a muestra la tasa de mortalidad global en Estados Unidos (o tasa de muertes), en muertes por 1 000 personas, para diferentes grupos de edad. Observe que hay un elevado riesgo de muerte cerca del nacimiento, después las tasas caen a niveles muy bajos. Alrededor de los 15 años, las tasas de mortalidad empiezan a crecer gradualmente.

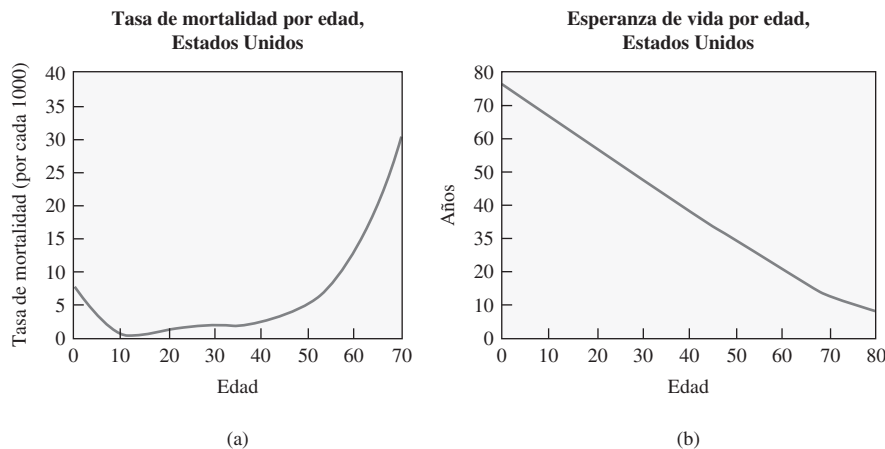


Figura 6.12 (a) La tasa global de mortalidad en Estados Unidos (muertes por cada 1 000 personas) para diferentes edades. (b) Esperanza de vida para diferentes edades. *Fuente:* Centro Nacional para Estadísticas de Salud, Estados Unidos.

La figura 6.12b muestra la **esperanza de vida** de estadounidenses de diferentes edades, definida como el número de años promedio que una persona de una edad dada puede esperar vivir. Como esperaríamos, la esperanza de vida es más alta para gente más joven ya que, en promedio, les queda más tiempo por vivir. Al nacer, la esperanza de vida de los estadounidenses actualmente es de alrededor de 78 años (75 años para hombres y 81 años para mujeres).

Definición

La **esperanza de vida** es el número de años que en promedio una persona con una edad dada ahora puede esperar vivir.

La sutileza en la interpretación de esperanza de vida proviene de los cambios en la ciencia médica y la salud pública. Las esperanzas de vida se calculan estudiando las tasas de mortalidad *actuales*. Por ejemplo, cuando decimos que la esperanza de vida actual de niños al nacer es 78 años, queremos decir que en promedio los niños vivirán hasta la edad de 78 años, *si no hay cambios futuros* en la ciencia médica o en salud pública. Así, mientras la esperanza de vida proporciona una medida útil de salud global actual, no debe considerarse como una *predicción* de tiempo de vida en el futuro.

De hecho, a consecuencia de los avances tanto de la ciencia médica como de la salud pública, las esperanzas de vida han aumentado significativamente en el siglo pasado, elevándose alrededor de 60% (figura 6.13). Si esta tendencia continúa, los infantes de hoy es probable que en promedio vivan más de 78 años.

A propósito...

Las esperanzas de vida varían ampliamente en el mundo, y la esperanza de vida en Estados Unidos (78 años) está a la zaga de muchos otros países desarrollados incluyendo Japón (81 años), Australia (80.5 años) y Canadá (80.2 años). Las menores esperanzas de vida se encuentran en los países asolados por enfermedades y por la guerra del África subsahariana, tal como Angola (38 años) y Sierra Leona (40 años).

Los reportes de mi muerte son muy exagerados.

—Mark Twain, desde Londres en un cable enviado a Associated Press

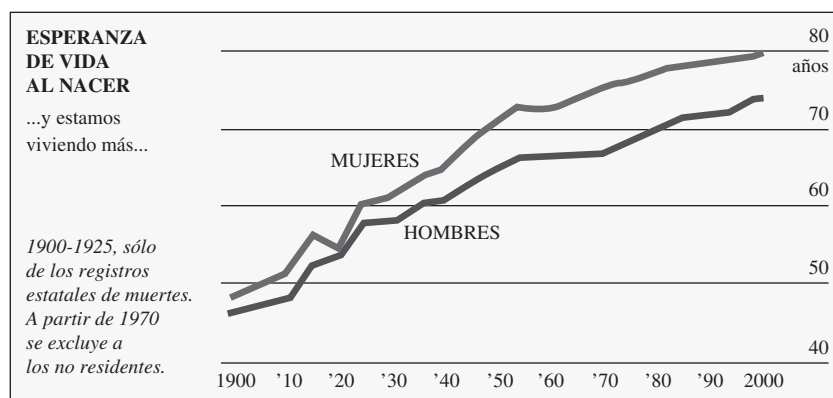


Figura 6.13 Cambios en la esperanza de vida en Estados Unidos durante el siglo xx.

Fuente: *New York Times* y Centro Nacional para Estadísticas de Ciencias de la Salud.

UN MOMENTO DE REFLEXIÓN

Usando la figura 6.13, compare las esperanzas de vida de hombres y mujeres. Analice brevemente estas diferencias. ¿Tienen alguna implicación para la política social? ¿Para tarifas de seguros? Explique.

EJEMPLO 4 Esperanzas de vida

Usando la figura 6.12b determine la esperanza de vida de una persona de 20 años de edad y de una persona de 60 años de edad. ¿Los números son consistentes? Explique.

Solución La gráfica muestra que la esperanza de vida a la edad de 20 es alrededor de 58 años y a la edad de 60 es alrededor de 21 años. Esto significa que una persona de 20 años de edad puede esperar vivir, en promedio, alrededor de 58 años más, para cumplir 78. Un promedio de 60 años puede esperar vivir alrededor de 21 años más para llegar a 81.

En principio podría parecer extraño que alguien de 60 años tenga un promedio de vida mayor que alguien de 20 años (81 años frente a 78 años). Pero recuerde que las esperanzas de vida tienen como base datos *actuales*. Si no hubiese cambios en medicina o salud pública, una persona de 60 años tendría una mayor probabilidad de llegar a 81 que uno de 20 años, ya que él o ella ya habrán llegado a 60. Sin embargo, si la medicina y la salud pública continúan mejorando, una persona de 20 años ahora podría llegar a una edad más avanzada que una persona que actualmente tenga 60 años.

ESTUDIO DE CASO

Esperanza de vida y seguridad social

Debido al cambio de la conformación de edades en la población de Estados Unidos, el número de jubilados que cobrarían beneficios se espera que sean mucho mayores en el futuro, mientras que el número de asalariados que pagan impuestos de seguridad social, se espera que permanezcan relativamente constantes. Como resultado, uno de los mayores retos para el futuro de la seguridad social es determinar una forma de asegurar el dinero suficiente para pagar los beneficios de jubilados.

Las proyecciones actuales muestran que, sin cambios significativos, el programa de seguridad social estará en quiebra e incapaz de pagar los beneficios completos después de 2040. (Esto puede sonar muy lejano pero, en ese momento, los estudiantes universitarios de ahora estarán en sus principales años de generación de ingresos). Las autoridades de seguridad social han propuesto diferentes maneras de resolver este problema, incluyendo cambios en los montos de beneficios pagados, aumento en las tasas de impuestos de seguridad social, cambios en la edad de jubilación y programas de seguridad social parcial o completamente privatizados. Cada una de estas propuestas enfrenta obstáculos políticos. Pero, además, todas estas propuestas están basadas en la suposición acerca de la esperanza de vida futura. Si no hay cambios

correspondientes en la edad de retiro, vidas más largas significan más años de beneficios de seguridad social.

Más específico, propuestas recientes de autoridades de seguridad social han supuesto que la esperanza de vida al nacer se elevará a 79.3 en 2030 y a 81.5 en 2070 (para ambos sexos). Pero estos números pueden estar lejos de ser demasiado pesimistas. Por ejemplo, las proyecciones de seguridad social suponen que las mujeres estadounidenses no llegarán a una esperanza de vida de 82 años hasta 2033, pero las mujeres en muchos países de Europa y de Asia ya han alcanzado esta esperanza de vida. ¿En realidad le tomará a Estados Unidos décadas alcanzarlas? Un panel de expertos sugirió que estimaciones más realistas de esperanza de vida futura sería 81 en 2030 y 85 en 2070, lo que significa casi cuatro años más de beneficios de seguridad social para la persona promedio en 2070. Y, si la ciencia médica logra dar un gran paso que permita a las personas vivir más, la crisis de la seguridad social podría ser mucho peor.

UN MOMENTO DE REFLEXIÓN

De acuerdo con algunos biólogos, existe una buena posibilidad de que los avances del siglo XXI en la ciencia médica permitan que la gente viva 100 años o más. ¿Cómo afectaría a programas como los de seguridad social? ¿Qué otros efectos esperaría tener en la sociedad? Globalmente, ¿considera que grandes aumentos en la esperanza de vida serían buenos o malos para la sociedad? Defienda su opinión.

Sección 6.4 Ejercicios

Alfabetización estadística y pensamiento crítico

- Tasa de natalidad.** La tasa de natalidad actual en Estados Unidos está dada como 14.1 por cada 1000 personas. Cuando se compara el crecimiento de población en países diferentes, ¿por qué es mejor utilizar *tasas* en lugar de los números reales de nacimientos?
- Estadísticas vitales.** ¿Qué son las estadísticas vitales?
- Esperanza de vida.** ¿Qué es la esperanza de vida? ¿Una persona de 30 años de edad tiene la misma esperanza de vida que una persona de 20 años de edad? ¿Por qué sí o por qué no?
- Esperanza de vida.** Con base en datos recientes, una persona de 20 años de edad en Estados Unidos tiene una esperanza de vida de 57.8 años. ¿Qué significa esto?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Accidente aéreo.** Los accidentes aéreos reciben mucha atención en los medios, incluso si el accidente involucra a una o dos personas. Puesto que a los accidentes aéreos se les da tanta atención, el riesgo de morir en uno es mayor que el riesgo de morir en un accidente automovilístico.
- Esperanza de vida.** Un funcionario de salud pública indica que su esperanza de vida disminuye con su edad.

7. Edad esperada de muerte. Conforme se hace más viejo, su edad esperada de morir aumenta.

8. Riesgo de morir. En un año reciente los números totales de muertes en Estados Unidos debidas a accidentes y neumonía fueron aproximadamente iguales. Por tanto, los riesgos de morir por accidente y neumonía por cada 100000 personas son aproximadamente iguales.

Tasas de accidentes aéreos mortales en vuelos comerciales. Para los ejercicios del 9 al 12 utilice la tabla siguiente, que resume datos en vuelos comerciales en Estados Unidos para tres años separados.

Año	Salidas (miles)	Accidentes fatales	Millas de pasajeros (miles de millones)	Pasajeros (millones)
1995	8062	168	540.7	547.8
2000	9035	92	692.8	666.2
2004	11182	14	731.9	697.8

- Para cada uno de los tres años determine la tasa de accidentes fatales en muertes por cada 1000 salidas. Con base en esas tasas, ¿cuál año fue el más seguro? ¿Por qué?
- Para cada uno de los tres años, determine la tasa de accidentes fatales en muertes por cada 10000 millones de millas de pasajeros. Con base en esas tasas, ¿cuál año fue el más seguro? ¿Por qué?
- Para cada uno de los tres años determine la tasa de accidentes fatales en muertes por cada 10 millones de pasajeros. Con base en esas tasas, ¿cuál año fue el más seguro? ¿Por qué?

12. Para el año 2004 determine la tasa de accidentes fatales en muertes por milla de pasajero. ¿Por qué no reportamos la tasa de fatalidad en unidades de muertes por milla de pasajero?

Tabla de vida. Para los ejercicios 13 al 16 utilice los datos en la tabla siguiente para personas en Estados Unidos entre edades de 16 y 21 años.

Intervalo de edad	Probabilidad de morir durante el intervalo	Número de sobrevivientes al inicio del intervalo	Número de muertes durante el intervalo	Esperanza de vida restante (desde el inicio del intervalo)
16-17	0.000607	98 943	60	61.7
17-18	0.000706	98 883	70	60.7
18-19	0.000780	98 814	77	59.7
19-20	0.000833	98 736	82	58.8
20-21	0.000888	98 654	88	57.8

13. **Tiempo de vida esperado.** ¿Cuántos años se espera que viva una persona de 19 años seleccionada aleatoriamente después de su cumpleaños 19?
14. **Tiempo de vida esperado.** ¿Cuántos años se espera que viva una persona de 17 años seleccionada aleatoriamente después de su cumpleaños 17?
15. **Tasa de mortalidad.** Las compañías de seguros deben estar muy al tanto de las tasas de muerte. Antes de emitir una póliza de seguro de vida para una persona de 19 años, la compañía necesita conocer la tasa de mortalidad para ese grupo de edad. Determine la tasa de mortalidad por cada 10 000 para personas durante sus 19 años.
16. **Tasa de mortalidad.** Determine la tasa de mortalidad por cada 10 000 personas durante sus 16 años.
17. **Tasas de natalidad alta/baja en Estados Unidos.** Las *tasas* de natalidad más alta y más baja en Estados Unidos en 2003 fueron en Utah y Maine, respectivamente. Utah reportó 49 870 nacimientos con una población de alrededor de 2.4 millones de personas. Maine reportó 13 861 nacimientos con una población de alrededor de 1.3 millones de personas. Utilice estos datos para responder las preguntas siguientes.
- ¿Cuántas personas nacieron por día en Utah? ¿Cuántas en Maine?
 - ¿Cuál fue la tasa de natalidad en Utah en nacimientos por cada 1 000 personas? ¿Cuál fue en Maine?
18. **Números de muertes alta/baja.** En un año reciente, hubo 235 000 muertes en California, el número más grande en Estados Unidos. El estado con el número más bajo de muertes fue Alaska, con 3 000 decesos. Las poblaciones de California y Alaska fueron aproximadamente 35 463 000 y 648 000, respectivamente.

- Calcule las *tasas* de mortalidad para California y Alaska en muertes por 1 000.
- Con base en el hecho que California y Alaska tuvieron los números de muertes mayor y menor, respectivamente, ¿se deduce que California y Alaska tuvieron las *tasas* más alta y más baja de mortalidad? ¿Por qué sí o por qué no?

19. **Tasas de natalidad y mortalidad en Estados Unidos.** En 2006 la población de Estados Unidos alcanzó 300 millones. La tasa de natalidad global fue de 14.1 nacimientos por cada 1 000 y la tasa de mortalidad global se estimó en 8.4 muertes por cada 1 000.

- Aproximadamente, ¿cuántos nacimientos hubo en Estados Unidos?
- ¿Alrededor de cuántas muertes hubo en Estados Unidos?
- Con base sólo en los nacimientos y las muertes (sin contar inmigración y emigración), alrededor de cuánto se elevó la población de Estados Unidos durante 2006?
- Ignorando la inmigración y la emigración, ¿cuál es la tasa de crecimiento de la población en 2006 en Estados Unidos? ¿Cuál es la tasa de crecimiento de población expresada como porcentaje?

20. **Tasas de natalidad y mortalidad en China.** Las estimaciones siguientes de 2006 son para China: población = 1 313 973 713; tasa de natalidad = 13.25 por cada 1 000; tasa de mortalidad = 6.97 por cada 1 000.

- Aproximadamente, ¿cuántos nacimientos hubo en China?
- ¿Alrededor de cuántas muertes hubo en China?
- Con base sólo en los nacimientos y las muertes (sin tomar en cuenta inmigración y emigración), alrededor de cuánto aumentó la población de China durante 2006?
- Ignorando inmigración y emigración, ¿cuál fue la tasa de crecimiento de la población en China en 2006? ¿Cuál es la tasa de crecimiento expresada como porcentaje?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 6 en www.aw.com/bbt.

21. **Esperanza de vida, Estados Unidos frente al mundo.** Puede encontrar una gran cantidad de información en la web acerca de esperanzas de vida en el mundo. ¿Cómo se compara la esperanza de vida de Estados Unidos con la de otros países desarrollados? ¿Qué podría explicar la diferencia que ve? Con base en sus hallazgos analice las posibles implicaciones para las políticas sociales o gubernamentales en Estados Unidos.
22. **Esperanzas de vida de hombres y mujeres.** Encuentre datos acerca de cómo y por qué las esperanzas de vida de hombres y mujeres están cambiando con el tiempo. ¿Por qué las mujeres tienen mayores esperanza de vida que los

hombres? En el futuro, ¿las esperanzas de vida de los hombres alcanzarán las esperanzas de vida de las mujeres? Resuma sus hallazgos con un reporte breve.

23. **Caso de estudio, Uganda.** Encuentre datos que consideren las esperanzas de vida en Uganda en las décadas recientes. Verá que la esperanza de vida en Uganda se ha elevado espectacularmente en la década pasada. ¿Qué explica este aumento y qué lecciones podría tener para mejorar la esperanza de vida en otros países africanos?
24. **Cálculos de esperanza de vida.** Usted encontrará muchas calculadoras de esperanza de vida en internet, trate con algunas de ellas. ¿Parecen dar resultados precisos o realistas? Explore las técnicas estadísticas que se utilizaron para hacer las tablas de esperanza de vida.
25. **Escala Richter para riesgo.** La Real Sociedad de Estadística ha propuesto un sistema de magnitudes de riesgo y factores de riesgo análogos a la escala Richter para medir terremotos. Vaya a internet para aprender cómo estas medidas de riesgo se definen y calculan. Con estas medidas, analice los riesgos de varias actividades y eventos.
26. **Comprensión del riesgo.** El libro *Against the Gods: The Remarkable Story of Risk* de Peter Bernstein (John Wiley, 1996) es un relato galardonado de la historia de la proba-

bilidad y la valuación del riesgo. Encuentre el libro en una biblioteca o en una librería (es una compra valiosa) e identifique un evento particular que cambió nuestra comprensión del riesgo. Escriba un ensayo de dos páginas tanto de la historia como de las consecuencias de este evento en particular.

EN LAS NOTICIAS

27. **Seguridad al viajar.** Encuentre una noticia reciente que analice algún aspecto de seguridad al viajar (tal como riesgo de accidentes en automóviles o aviones, la eficacia de asientos para niños en automóviles, o los efectos de manejar mientras se habla por un teléfono celular). Resuma cualquier estadística dada acerca de los riesgos y proporcione su opinión general considerando el tema de seguridad bajo discusión.
28. **Estadísticas vitales.** Encuentre una noticia que proporcione datos actuales sobre estadísticas vitales o esperanzas de vida. Resuma el reporte y las estadísticas, y analice cualquier implicación personal o social de los nuevos datos.

6.5 Combinación de probabilidades (sección complementaria)

Las ideas de probabilidad que hemos analizado en el capítulo hasta ahora serán suficientes para la mayoría del trabajo que haremos en este libro. Sin embargo, la probabilidad tiene muchas más aplicaciones, tanto en estadística como en otras áreas de la vida. En esta sección investigaremos unas ideas más de probabilidad y unas cuantas de sus muchas aplicaciones.

Probabilidades y

Suponga que tira dos dados y quiere conocer la probabilidad que *ambos* caigan en 4. Una manera para determinar la probabilidad es considerar el tiro de dos dados en sucesión como un *solo* tiro de dos dados. Entonces podemos encontrar la probabilidad usando el método *teórico* (vea la sección 6.2). Ya que “doble 4” es 1 de 36 posibles resultados, su probabilidad es $1/36$.

De manera alterna, podemos considerar los dos dados de modo individual. Para cada dado la probabilidad de un 4 es $1/6$. Determinamos la probabilidad de que ambos dados muestren 4 multiplicando las probabilidades individuales:

$$P(\text{doble } 4) = P(4) \times P(4) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

Por cualquier método, la probabilidad de obtener doble 4 es $1/36$. En general, llamamos a la probabilidad de que ocurra el evento *A* y el evento *B* una **probabilidad de “y”** (o *probabilidad conjunta*).

La ventaja de la técnica de multiplicación es que puede extenderse con facilidad a situaciones que incluyan más de dos eventos. Por ejemplo, podría necesitar determinar la probabilidad de obtener 10 caras en 10 lanzamientos de una moneda, o de tener un bebé y obtener un

La suerte sólo favorece a la mente preparada

—Louis Pasteur,
científico del siglo XIX

Apropósito...

Si los eventos son independientes no siempre es obvio. Considere un conjunto de tiros libres en básquetbol realizados por una sola jugadora. Algunas personas argumentan que la probabilidad de encestar en cada tiro libre es la misma. Otros sugieren que el impacto psicológico de éxito o fracaso en un tiro libre afecta las posibilidades de la jugadora en el siguiente tiro. A pesar de muchos estudios estadísticos de tiros libres en básquetbol, nadie ha establecido claramente cuál punto de vista es correcto.



aumento de salario en el mismo año. Sin embargo, existe una diferencia importante que debe hacerse cuando se trabaja con probabilidades y. Debemos distinguir entre eventos que son *independientes* y eventos que son *dependientes* entre ellos. Analicemos cada caso.

Eventos independientes

El tiro repetido de un solo dado produce **eventos independientes** ya que el resultado de un tiro no afecta las probabilidades de los otros tiros. De manera análoga, lanzamientos de una moneda son independientes. La decisión de si eventos son independientes es importante, pero la respuesta no siempre es obvia. Por ejemplo, al analizar una sucesión de tiros libres de un jugador de básquetbol, ¿debemos suponer que un tiro libre es independiente de los otros? Siempre que los eventos *sean* independientes, podemos calcular la probabilidad y de dos o más eventos multiplicando.

Probabilidad de “y” para eventos independientes

Dos eventos son **independientes** si el resultado de un evento no afecta la probabilidad del otro evento. Considere dos eventos independientes A y B con probabilidades $P(A)$ y $P(B)$. La probabilidad de que A y B ocurran juntos es

$$P(A \text{ y } B) = P(A) \times P(B)$$

Este principio puede extenderse a cualquier número de eventos independientes. Por ejemplo, la probabilidad de A , B y un tercer evento independiente C es

$$P(A \text{ y } B \text{ y } C) = P(A) \times P(B) \times P(C)$$

EJEMPLO 1 Tres monedas

Suponga que lanza tres monedas justas. ¿Cuál es la probabilidad de obtener tres cruces?

Solución Puesto que los lanzamientos de monedas son independientes, multiplicamos la probabilidad de cruz en cada moneda individual:

$$P(3 \text{ cruces}) = \underbrace{P(\text{cruz})}_{\text{moneda 1}} \times \underbrace{P(\text{cruz})}_{\text{moneda 2}} \times \underbrace{P(\text{cruz})}_{\text{moneda 3}} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

La probabilidad de que tres monedas lanzadas caigan en cruz es $1/8$ (que se determinó en el ejemplo 8 de la sección 6.2 con mucho más trabajo).

Eventos dependientes

Un lote de 15 tarjetas de memoria tiene 5 tarjetas defectuosas. Si selecciona una tarjeta al azar del lote, la probabilidad de obtener una defectuosa es $5/15$. Ahora suponga que selecciona una defectuosa en la primera selección y la mete a su bolsa. ¿Cuál es la probabilidad de obtener una defectuosa en la segunda selección?

Puesto que ha quitado una tarjeta defectuosa del lote, éste ahora sólo tiene 14 tarjetas, de las cuales 4 son defectuosas. Así, la probabilidad de obtener una tarjeta defectuosa en la segunda extracción es $4/14$. Esta probabilidad es menor que la probabilidad $5/15$ en la primera selección, ya que la primera selección cambió el contenido del lote. Puesto que el resultado del primer evento afecta la probabilidad del segundo evento, éstos son **eventos dependientes**.

El cálculo de la probabilidad de eventos dependientes sigue involucrando la multiplicación de probabilidades individuales, pero debemos tomar en cuenta cómo los eventos previos afectan los eventos subsecuentes. En el caso del lote de tarjetas de memoria, encontramos la probabilidad de obtener dos tarjetas defectuosas seguidas multiplicando la probabilidad $5/15$, para la primera selección, por la probabilidad $4/14$, de la segunda selección.

$$P(2 \text{ defectuosas}) = \underbrace{P(\text{defectuosa})}_{\text{primera selección}} \times \underbrace{P(\text{defectuosa})}_{\text{segunda selección, si la primera fue defectuosa}} = \frac{5}{15} \times \frac{4}{14} = 0.0952$$

La probabilidad de sacar dos tarjetas defectuosas seguidas es 0.0952, que es un poco menor que $(5/15) \times (5/15) = 0.111$, la probabilidad que obtenemos si reemplazamos la primera carta antes de la segunda selección.

Probabilidad de “y” para eventos dependientes

Dos eventos son **dependientes** si el resultado de un evento afecta la probabilidad del otro evento. La probabilidad de que eventos dependientes ocurran juntos es

$$P(A \text{ y } B) = P(A) \times P(B \text{ dado } A)$$

donde $P(B \text{ dado } A)$ significa la probabilidad del evento B dada la ocurrencia del evento A .

Este principio puede extenderse a cualquier número de eventos individuales. Por ejemplo, la probabilidad de los eventos dependientes A , B y C es

$$P(A \text{ y } B \text{ y } C) = P(A) \times P(B \text{ dado } A) \times P(C \text{ dados } A \text{ y } B)$$

NOTA TÉCNICA

$P(B \text{ dado } A)$ se denomina *probabilidad condicional*. En algunos libros se denota $P(B|A)$.

EJEMPLO 2 Jugando BINGO

El juego de BINGO incluye sacar al azar, de una bolsa, botones rotulados, sin reemplazarlos. Hay 75 botones, 15 para cada una de las letras B, I, N, G y O. ¿Cuál es la probabilidad de sacar dos botones B en las primeras dos selecciones?

Solución BINGO implica eventos dependientes porque quitar un botón cambia el contenido de la bolsa. La probabilidad de sacar una B en la primera extracción es 15/75. Si esto ocurre, 74 botones quedan en la bolsa, de los cuales 14 son B. Por tanto, la probabilidad de sacar un botón B en la segunda extracción es 14/74. La probabilidad de sacar dos botones B en las primeras dos selecciones es

$$P(B \text{ y } B) = \underbrace{P(B)}_{\text{primera extracción}} \times \underbrace{P(B)}_{\text{segunda extracción, dada una B en la primera extracción}} = \frac{15}{75} \times \frac{14}{74} = 0.0378$$

UN MOMENTO DE REFLEXIÓN

Sin hacer cálculos compare la probabilidad de sacar tres corazones seguidos cuando la carta es reemplazada y el mazo se baraja después de cada extracción con la probabilidad de sacar tres corazones seguidos sin reemplazar la carta. ¿Cuál probabilidad es mayor y por qué?

EJEMPLO 3 ¿Cuándo podemos tratar eventos dependientes como eventos independientes?

Una organización de encuestas tiene una lista de 1 000 personas para una encuesta telefónica. Los encuestadores saben que 433 personas de las 1 000 son miembros del Partido Demócrata. Suponiendo que una persona no puede ser llamada más de una vez, ¿cuál es la probabilidad de que las primeras dos personas llamadas sean miembros del Partido Democrático?

Solución Este problema implica una probabilidad y para eventos dependientes. Una vez que una persona es llamada, esa persona no puede ser llamada otra vez. La probabilidad de llamar a un miembro del Partido Demócrata en la primera llamada es 433/1 000. Con esa persona eliminada del grupo de llamadas, la probabilidad de llamar a un miembro del Partido Demó-

La importancia de la probabilidad sólo puede ser derivada del juicio que sea racional para ser guiada por él en la acción.

—John Maynard Keynes

crata en la segunda llamada es 432/999. Por tanto, la probabilidad de llamar a dos miembros del Partido Demócrata en las primeras dos llamadas es

$$\frac{433}{1000} \times \frac{432}{999} = 0.1872$$

Si tratamos las dos llamadas como eventos independientes, la probabilidad de llamar a un miembro del Partido Demócrata es 433/1000 en ambas llamadas. Entonces la probabilidad de llamar a dos miembros de este partido es

$$\frac{433}{1000} \times \frac{433}{1000} = 0.1875$$

que es casi idéntico al resultado de suponer eventos dependientes. En general, si relativamente pocos elementos o personas son seleccionados de un grupo grande (en este caso, 2 personas de 1000), entonces los eventos dependientes pueden tratarse como eventos independientes con muy poco error. Una directriz común es que la independencia puede suponerse cuando el tamaño de la muestra es menor a 5% del tamaño de la población. Esta práctica es comúnmente usada por organizaciones de encuestas.

Apropósito...

A finales de la década de los sesenta, el famoso doctor Benjamín Spock fue declarado culpable, por un jurado compuesto únicamente por hombres, por alentar la resistencia al reclutamiento durante la guerra de Vietnam. Su defensa argumentó que un jurado con mujeres habría sido más compasivo.



EJEMPLO 4 Selección de jurado

Un jurado de nueve personas se selecciona al azar de un conjunto grande de personas que tiene igual número de hombres y mujeres. ¿Cuál es la probabilidad de seleccionar un jurado compuesto sólo por hombres?

Solución Cuando un miembro del jurado se selecciona, éste se elimina del grupo. Sin embargo, puesto que estamos seleccionando un número pequeño de miembros para el jurado de un grupo grande, podemos tratar la selección de éstos como eventos independientes. Suponiendo igual número de hombres y de mujeres disponibles, la probabilidad de seleccionar un miembro del jurado masculino es $P(\text{hombre}) = 0.5$. Así, la probabilidad de seleccionar 9 hombres como miembros del jurado es

$$P(9 \text{ hombres}) = \underbrace{0.5 \times 0.5 \times \dots \times 0.5}_{9 \text{ veces}} = 0.5^9 = 0.00195$$

La probabilidad de seleccionar un jurado compuesto únicamente por hombres elegidos al azar es 0.00195, o aproximadamente 2 en 1000. (Nota: expresiones de la forma a^b , tal como 0.5^9 , pueden evaluarse en muchas calculadoras usando la tecla marcada x^y o y^x o $>$).

Probabilidades *cualquiera/o*

Hasta ahora hemos considerado la probabilidad de que ocurra un evento en un primer ensayo y otro evento ocurra en un segundo ensayo. Suponga que queremos conocer la probabilidad de que, cuando un ensayo es realizado, *cualquiera* de dos eventos ocurra. En ese caso estamos buscando una **probabilidad de cualquiera/o** tal como la probabilidad de tener un hijo con ojos azules *o* con ojos verdes, o bien, la probabilidad de perder su casa por un incendio *o* por un huracán.

Eventos que no se traslapan

Una moneda puede caer en cara *o* en cruz, pero no puede caer en cara y en cruz al mismo tiempo. Cuando no es posible que dos eventos puedan ocurrir al mismo tiempo, se dice que son **eventos que no se traslapan** (o *mutuamente excluyentes*). Podemos representar eventos que no se traslapan con un diagrama de Venn en el que cada círculo representa un evento. Si los círculos no se traslapan, significa que los eventos correspondientes no pueden ocurrir juntos. Por ejemplo, mostramos las posibilidades de cara y cruz en el lanzamiento de una moneda como dos círculos que no se traslapan, ya que una moneda no puede caer al mismo tiempo en cara y en cruz (figura 6.14).

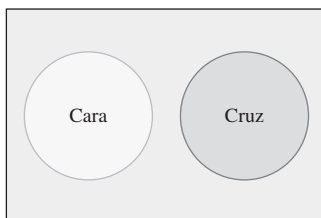


Figura 6.14 Diagrama de Venn para eventos que no se traslapan.

Suponga que tiramos un dado una vez y queremos encontrar la probabilidad de lanzar un 1 o un 2. Puesto que hay seis resultados igualmente probables y puesto que dos de esos resultados corresponden a los eventos en cuestión, podemos utilizar el método teórico para concluir que la probabilidad de tirar un 1 o un 2 es $P(1 \text{ o } 2) = 2/6 = 1/3$. De manera alterna, podemos determinar esta probabilidad sumando las probabilidades individuales $P(1) = 1/6$ y $P(2) = 1/6$. En general, encontramos la probabilidad de *cualquiera/o* de dos eventos que no se traslapan sumando las probabilidades individuales.

Probabilidad de *cualquiera/o* para eventos que no se traslapan

Dos eventos son **no traslapados**, si ellos no pueden ocurrir al mismo tiempo. Si A y B son eventos que no se traslapan, la probabilidad de que cualquiera A o B ocurra es

$$P(A \text{ o } B) = P(A) + P(B)$$

Este principio se puede extender para cualquier número de eventos que no se traslapan. Por ejemplo, la probabilidad de que cualquiera de los eventos A , B o C ocurra es

$$P(A \text{ o } B \text{ o } C) = P(A) + P(B) + P(C)$$

siempre que A , B y C sean eventos que no se traslapan.

EJEMPLO 5 Probabilidad *cualquiera/o* para un dado

Suponga que tira un solo dado. ¿Cuál es la probabilidad de obtener un número par?

Solución Los resultados pares 2, 4 y 6 no se traslapan ya que un dado sólo puede arrojar un resultado. Sabemos que $P(2) = 1/6$, $P(4) = 1/6$ y $P(6) = 1/6$. Por tanto, la probabilidad combinada es

$$P(2 \text{ o } 4 \text{ o } 6) = P(2) + P(4) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

La probabilidad de obtener un número par es $1/2$.

Eventos que se traslapan

Para mejorar el turismo entre Francia y Estados Unidos, los dos gobiernos forman un comité que consiste de 20 personas, 2 hombres estadounidenses, 4 hombres franceses, 6 mujeres estadounidenses y 8 mujeres francesas (como se muestra en la tabla 6.9). Si se encuentra con una de estas personas al azar, ¿cuál es la probabilidad de que la persona será *mujer o* una persona de Francia?

Tabla 6.9 Comité de turismo		
	Hombres	Mujeres
Estadounidense	2	6
Francés	4	8

Doce de las 20 personas son francesas, por lo que la probabilidad de encontrarse con una persona francesa es $12/20$. De forma análoga, 14 de las 20 personas son mujeres, por lo que la probabilidad de encontrarse con una mujer es $14/20$. La suma de estas dos probabilidades es

$$\frac{12}{20} + \frac{14}{20} = \frac{26}{20}$$

Ésta no puede ser la probabilidad correcta de encontrarse con una mujer o con una persona de Francia, ya que las probabilidades no pueden ser mayores a 1. El diagrama de Venn en la figura 6.15 muestra por qué, en esta situación, la simple suma es incorrecta. El círculo de la izquierda contiene a las 12 personas francesas, el círculo de la derecha contiene a las 14 mujeres y los hombres estadounidenses no aparecen en los círculos. Vemos que existen 18 personas que son francesas o mujeres (o ambas). Puesto que el total de personas es 20, la probabilidad de encontrarse con una persona que es *de* Francia o mujer es $18/20 = 9/10$.

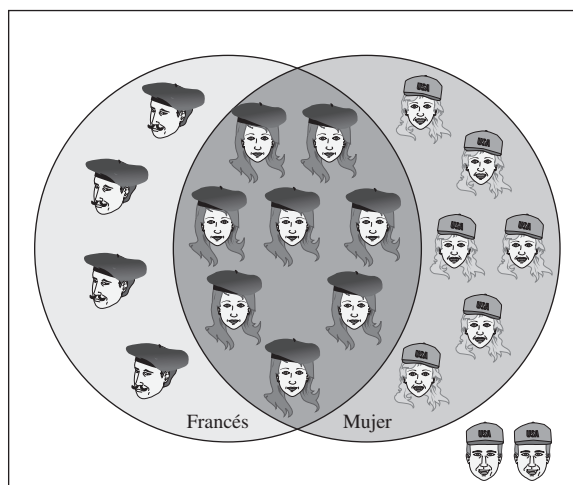


Figura 6.15 Diagrama de Venn para eventos que se traslapan.

Como muestra el diagrama de Venn, la suma simple fue incorrecta ya que la región en que los círculos se traslapan contiene 8 personas que *a la vez* son de Francia y mujeres. Si sumamos las dos probabilidades individuales, estas 8 personas se cuentan dos veces: una vez como mujer y una vez como francesa. La probabilidad de encontrarse con una de estas mujeres francesas es $8/20$. Podemos corregir el error de conteo doble restando esta probabilidad. Así, la probabilidad de encontrarse con una persona que es de Francia o es mujer es

$$P(\text{mujer o de Francia}) = \underbrace{\frac{14}{20}}_{\text{probabilidad mujer}} + \underbrace{\frac{12}{20}}_{\text{probabilidad persona de Francia}} - \underbrace{\frac{8}{20}}_{\text{probabilidad mujer francesa}} = \frac{18}{20} = \frac{9}{10}$$

que coincide con el resultado encontrado mediante conteo.

Decimos que encontrarse con una mujer y encontrarse con una persona francesa son eventos **traslapados** (*no mutuamente no excluyentes*) ya que ambos pueden ocurrir al mismo tiempo. Al generalizar el procedimiento usado en este ejemplo, encontramos la regla siguiente.

Probabilidad de *cualquiera/o* para eventos que se traslapan

Dos eventos A y B están **traslapados**, si pueden ocurrir juntos. La probabilidad de que ocurra cualquiera de A o B es

$$P(A \circ B) = P(A) + P(B) - P(A \text{ y } B)$$

El último término, $P(A \text{ y } B)$, corrige el conteo doble de los eventos en los que A y B ocurren juntos. Observe que no es necesario utilizar esta fórmula. La probabilidad correcta siempre puede determinarse mediante un conteo cuidadoso y evitar el doble conteo.

UN MOMENTO DE REFLEXIÓN

¿Los eventos de haber nacido un miércoles o haber nacido en Las Vegas se traslapan?
¿Los eventos de nacer en miércoles o nacer en marzo se traslapan? ¿Los eventos de nacer en miércoles o nacer en viernes se traslapan? Explique.

EJEMPLO 6 Minorías y pobreza

Pine Creek es un poblado “promedio” estadounidense: de sus 2350 ciudadanos, 1950 son blancos, de los que 11%, o 215 personas, viven debajo del nivel de pobreza. De los 400 ciudadanos de minorías, 28%, o 112 personas, viven debajo del nivel de pobreza. Si usted visita Pine Creek, ¿cuál es la probabilidad de encontrarse (al azar) con una persona que es de una minoría o vive debajo del nivel de pobreza? (Este ejemplo es hipotético, pero los porcentajes son consistentes con la demografía nacional de Estados Unidos).

Solución Encontrarse con un ciudadano de la minoría y encontrarse con una persona que vive en pobreza son eventos que se traslapan. Es útil construir una pequeña tabla como la 6.10 que muestra cuántos ciudadanos hay en cada una de las cuatro categorías.

Tabla 6.10 Ciudadanos en Pine Creek		
	En pobreza	Arriba de la pobreza
Blanco	215	1735
Minoría	112	288

Debe verificar que las cifras en la tabla sean consistentes con la información dada y que el total en las cuatro categorías es 2350. Puesto que hay 400 ciudadanos de minorías, la probabilidad de encontrarse (aleatoriamente) a un ciudadano de minoría es $400/2350 = 0.170$. Puesto que hay $215 + 112 = 327$ personas que viven en pobreza, la probabilidad de encontrarse con un ciudadano en pobreza es $327/2350 = 0.139$. La probabilidad de encontrarse a una persona que es *a la vez* un ciudadano de minoría y una persona que vive en pobreza es $112/2350 = 0.0477$. De acuerdo con la regla para eventos que se traslapan, la probabilidad de encontrarse con un ciudadano o una persona que vive en pobreza es

$$P(\text{minoría o pobreza}) = 0.170 + 0.139 - 0.0477 = 0.261$$

La probabilidad de encontrarse con un ciudadano que es de una minoría o una persona que vive debajo del nivel de pobreza es alrededor de 1 en 4. Note la importancia de restar el término que corresponde a encontrarse con una persona que *a la vez* es un ciudadano de minoría y una persona que vive en pobreza.

Resumen

La tabla 6.11 proporciona un resumen de la fórmula que hemos utilizado en las probabilidades combinadas.

Tabla 6.11 Resumen de probabilidades combinadas			
Probabilidad y: eventos independientes	Probabilidad y: eventos dependientes	Probabilidad cualquiera/o: eventos que no se traslapan	Probabilidad cualquiera/o: eventos que se traslapan
$P(A \text{ y } B) =$	$P(A \text{ y } B) =$	$P(A \text{ o } B) =$	$P(A \text{ o } B) =$
$P(A) \times P(B)$	$P(A) \times P(B \text{ dado } A)$	$P(A) + P(B)$	$P(A) + P(B) - P(A \text{ y } B)$

Sección 6.5 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Independencia.** Con sus palabras indique el significado de que dos eventos sean independientes.
- 2. Eventos no traslapados.** Con sus palabras indique lo que significa que dos eventos no se traslapen.
- 3. Muestreo con reemplazo.** El profesor en una clase de 25 estudiantes selecciona de manera aleatoria a un estudiante y luego selecciona a otro estudiante al azar. Si los 25 estudiantes están disponibles para la segunda selección, ¿este muestreo es con reemplazo o sin reemplazo? ¿El segundo resultado es independiente del primero?
- 4. Opuestos.** Si dos eventos son opuestos, tal como A y $\text{no } A$, ¿deben ser eventos que no se traslapan? ¿Por qué sí o por qué no?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente. No todos los enunciados tienen respuestas definitivas; por lo que su explicación es más importante que la respuesta elegida.

- 5. Lotería.** Los números 5, 17, 18, 27, 36 y 41 se sacaron en la pasada lotería; no deben elegirse en la siguiente lotería ya que ahora es menos probable que ocurran.
- 6. Probabilidades combinadas.** La probabilidad de lanzar una moneda y obtener cara es 0.5. La probabilidad de seleccionar una carta roja cuando se saca una carta de un mazo barajado también es 0.5. Cuando lanza una moneda y saca una carta, la probabilidad de obtener cara o una carta roja es $0.5 + 0.5 = 1$.
- 7. Probabilidad cualquiera/o.** $P(A) = 0.5$ y $P(A \text{ o } B) = 0.4$.
- 8. Lotería.** Cuando se sacan los números de la lotería, la combinación 1, 2, 3, 4, 5 y 6 es menos probable que se saquen que otras combinaciones.

Conceptos y aplicaciones

- 9. Nacimientos.** Suponiendo que niños y niñas son igualmente probables y que el género de un niño no es afectado por el género de hermanos o hermanas, determine la probabilidad de que una pareja tenga tres niñas cuando tienen tres hijos.
- 10. Adivinando.** Un examen rápido consiste en una pregunta de falso/verdadero seguida de una pregunta de opción múltiple con cuatro posibles respuestas (a, b, c, d). Si ambas preguntas se responden con intentos al azar, determine la probabilidad de que ambas respuestas sean correctas. ¿Adivinar, parece ser una buena estrategia en este examen?

- 11. Contraseña.** La propietaria de una computadora nueva crea una contraseña con dos caracteres. Ella selecciona aleatoriamente una letra del alfabeto para el primer carácter y un dígito (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) para el segundo carácter. ¿Cuál es la probabilidad de que su contraseña sea K9? ¿Esta contraseña sería efectiva como disuasión contra alguien que trata de tener acceso a su computadora?
- 12. Vestirse de naranja para cazar.** Un estudio de lesiones en caza y el uso de ropa de caza naranja mostró que entre 123 cazadores lesionados cuando cometen error para la caza, 6 se vistieron de naranja (con base en datos de los Centros de Control de Enfermedades). Si un estudio posterior inicia con la selección aleatoria de cazadores de esta muestra de 123, determine la probabilidad de que los primeros dos cazadores seleccionados se vistiesen de naranja.
 - a. Suponga que el primer cazador es reemplazado antes de la selección del siguiente.
 - b. Suponga que el primer cazador no es reemplazado antes de que se seleccione el segundo.
 - c. En esta situación, ¿cuál selección tiene más sentido, con reemplazo o sin reemplazo? ¿Por qué?
- 13. Melodías en radio.** Un reproductor de MP3 está cargado con 60 selecciones musicales: 30 de rock, 15 de jazz y 15 de blues. El reproductor está en “reproducción aleatoria”, por lo que las selecciones se reproducen de manera aleatoria y pueden repetirse. ¿Cuál es la probabilidad de cada uno de los eventos siguientes?
 - a. Las primeras cuatro selecciones todas sean jazz.
 - b. Las primeras cinco selecciones todas sean blues.
 - c. La primera selección es jazz y la segunda es rock.
 - d. Entre las primeras cuatro selecciones, ninguna es rock.
 - e. La segunda selección es la misma que la primera.
- 14. Encuestas por teléfono.** Un encuestador telefónico tiene los nombres y números telefónicos de 45 votantes, 20 de los cuales son demócratas registrados y 25 republicanos registrados. Se hacen llamadas en un orden aleatorio. Suponga que quiere determinar la probabilidad de que las primeras dos llamadas son a republicanos.
 - a. ¿Estos eventos son dependientes o independientes? Explique.
 - b. Si los trata como eventos *dependientes*, ¿cuál es la probabilidad de que las primeras dos llamadas sean a republicanos?
 - c. Si los trata como eventos *independientes*, ¿cuál es la probabilidad de que las primeras llamadas sean a republicanos?
 - d. Compare los resultados de los incisos b y c.

Probabilidad y decisiones de la corte. Los datos en la tabla siguiente muestran los resultados al declararse culpable o no culpable en 1 028 casos criminales. Utilice los datos para responder los ejercicios 15-20.

	Declaración de culpable	Declaración inocente
Enviado a prisión	392	58
No enviado a prisión	564	14

Fuente: Brereton y Casper, "Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of the Criminal Courts", *Law and Society Review*, vol. 16, núm. 1.

15. ¿Cuál es la probabilidad de que un acusado seleccionado al azar se declare culpable o bien sea enviado a prisión?
16. ¿Cuál es la probabilidad de que un acusado seleccionado al azar se declare inocente o bien no sea enviado a prisión?
17. Si dos acusados diferentes se seleccionan al azar, ¿cuál es la probabilidad de que ambos hagan declaraciones de culpabilidad?
18. Si dos acusados diferentes se seleccionan al azar, ¿cuál es la probabilidad de que ambos fueran sentenciados a prisión?
19. Si un acusado es seleccionado al azar, ¿cuál es la probabilidad de que el acusado haga una declaración de culpable y fuese enviado a prisión?
20. Si un acusado es seleccionado al azar, ¿cuál es la probabilidad de que el acusado haga una declaración de culpable y no sea enviado a prisión?

Muertes de peatones. Para los ejercicios 21 a 26 utilice la tabla siguiente, la cual resume datos de muertes de peatones que fueron causadas por accidentes (con base en datos de la Administración Nacional de Seguridad de Tránsito en Autopistas).

		¿Peatón ebrio?	
		Sí	No
¿Conductor ebrio?	Sí	59	79
	No	266	581

21. Si una de las muertes de peatones se selecciona aleatoriamente, determine la probabilidad de que el peatón estuviese ebrio o el conductor estuviese ebrio.
22. Si una de las muertes de peatones se selecciona aleatoriamente, determine la probabilidad de que el peatón no estuviese ebrio o el conductor no estuviese ebrio.
23. Si una de las muertes de peatones se selecciona aleatoriamente, determine la probabilidad de que el peatón estuviese ebrio o el conductor no estuviese ebrio.
24. Si una de las muertes de peatones se selecciona aleatoriamente, determine la probabilidad de que el conductor estuviese ebrio o el peatón no estuviese ebrio.

25. Si se seleccionan dos muertes de peatones diferentes, determine la probabilidad de que ambos casos involucren a conductores ebrios.

26. Si se seleccionan dos muertes de peatones diferentes, determine la probabilidad de que en ambos casos los peatones estuviesen ebrios.

27. Pruebas de medicamento. Un medicamento contra una alergia se prueba dando éste a 120 personas y a 100 personas se les da un placebo. Un grupo de control consiste en 80 personas que no se les dio el tratamiento. El número de personas en cada grupo que mostró mejoras aparece en la tabla siguiente.

	Medicamento para la alergia	Placebo	Control	Total
Mejora	65	42	31	138
No hay mejora	55	58	49	162
Total	120	100	80	300

- a. ¿Cuál es la probabilidad de que a una persona seleccionada al azar en el estudio se le haya dado el medicamento y el placebo?
- b. ¿Cuál es la probabilidad de que una persona seleccionada al azar haya mejorado o no haya mejorado?
- c. ¿Cuál es la probabilidad de que a una persona seleccionada al azar se le haya dado el medicamento o haya mejorado?
- d. ¿Cuál es la probabilidad de que a una persona seleccionada al azar se le haya dado el medicamento y mejorado?

28. Filiaciones políticas. En una ciudad común (consistente con la demografía nacional), los adultos tienen las filiaciones políticas listadas en la tabla siguiente. Todas las cifras son porcentajes.

	Republicano	Demócrata	Independiente	Total
Hombres	17	15	18	50
Mujeres	14	20	16	50
Total	31	35	34	100

- a. ¿Cuál es la probabilidad de que una persona seleccionada al azar en la ciudad sea republicana o demócrata?
- b. ¿Cuál es la probabilidad de que una persona seleccionada al azar en la ciudad sea republicana o mujer?
- c. ¿Cuál es la probabilidad de que una persona seleccionada al azar en la ciudad sea independiente o bien hombre?
- d. ¿Cuál es la probabilidad de que una persona seleccionada al azar en la ciudad sea republicana y mujer?

29. Distribuciones de probabilidad y genética. Muchos rasgos son controlados por un gen dominante, **A**, y un gen recesivo, **a**. Suponga que dos padres llevan estos genes en la

proporción 3:1; esto es, la probabilidad que cualquiera de los padres dé el gen **A** es 0.75, y la probabilidad que cualquiera de los padres proporcione el gen **a** es 0.25. Suponga que los genes son seleccionados de cada padre de manera aleatoria. Para responder las preguntas siguientes, imagine 100 ensayos de “nacimientos”.

- ¿Cuál es la probabilidad de que un niño reciba un gen **A** de cada uno de los padres?
- ¿Cuál es la probabilidad de que un niño reciba un gen **A** de un padre y un gen **a** del otro? Observe que esto puede ocurrir de dos formas.
- ¿Cuál es la probabilidad de que un niño reciba un gen **a** de cada uno de los padres?
- Construya una tabla que muestre la distribución de probabilidad para todos los eventos.
- Si las combinaciones **AA** y **Aa** tienen como resultado el mismo rasgo dominante (digamos, cabello café) y **aa** tiene como resultado en el rasgo recesivo (digamos, cabello rubio), ¿cuál es la probabilidad de que un niño tenga el rasgo dominante?

30. BINGO. El juego de BINGO involucra sacar botones con letras y números al azar de un barril. Los B son del 1-15, los I son 16-30, los N son del 31-45, los G son del 46-60 y los O son del 61-75. Los botones no son reemplazados después de que se han seleccionado. ¿Cuál es la probabilidad de cada uno de los eventos siguientes en las selecciones iniciales?

- Sacar un botón B.
- Sacar dos botones B en secuencia.
- Sacar una B o una O.
- Sacar una B, luego una G y luego una N, en ese orden.
- Sacar algo distinto a N en cinco extracciones.

Problemas relativos a “al menos una vez”. Un problema común pide la probabilidad que un evento ocurra al menos una vez en un número dado de ensayos. Suponga que la probabilidad de un evento particular es p (por ejemplo, la probabilidad de sacar un corazón en un mazo de cartas es 0.25). Entonces la probabilidad que el evento ocurra al menos una vez en N ensayos es

$$1 - (1 - p)^N$$

Por ejemplo, la probabilidad de sacar al menos un corazón en 10 extracciones (con reemplazo) es

$$1 - (1 - 0.25)^{10} = 0.944$$

Utilice esta regla para resolver los ejercicios 31 y 32.

31. Las apuestas del Caballero de Mère. Se dice que la teoría de la probabilidad se inventó en el siglo XVII para explicar el juego a un noble conocido como Caballero de Mère.

- En su primer juego, el Caballero apostó a obtener al menos un 6 con cuatro tiros de un dado. Si jugó de manera repetida este juego, ¿debe esperar ganar?
- En el segundo juego, el Caballero apostó a obtener al menos un doble 6 con 24 tiros de dos dados. Si jugó de manera repetida, ¿éste es un juego en el que debe esperar ganar?

32. VIH en estudiantes universitarios. Suponga que 3% de los estudiantes en un colegio se sabe que son portadores del VIH.

- Si un estudiante tiene 6 parejas sexuales durante el curso de un año, ¿cuál es la probabilidad de que al menos una sea portadora del VIH?
- Si un estudiante tiene 12 parejas sexuales durante un año, ¿cuál es la probabilidad de que al menos una sea portadora de VIH?
- ¿Cuántas parejas necesitaría tener un estudiante antes de que la probabilidad de un encuentro VIH exceda 50%?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 6 en www.aw.com/bbt.

33. Simulación. Un problema clásico de probabilidad incluye a un rey que quiere aumentar la proporción de mujeres en su reino. Él promulga que después de que una madre dé a luz un varón, ella tiene prohibido tener más hijos. El rey razona que algunas familias tendrán un solo hijo, mientras que otras familias tendrán unas cuantas mujeres y un niño, así que la proporción de mujeres aumentará. Utilice el lanzamiento de una moneda para simular un reino que acata este decreto: “Después de que una madre da a luz a un varón, ella ya no podrá tener ningún otro hijo”. Si este decreto se sigue, ¿la proporción de mujeres aumentará?



EN LAS NOTICIAS

34. Un columnista del *New York Daily News* (Stephen Allensworth) proporciona sugerencias para seleccionar números en la lotería de Nueva York. Él defiende un sistema basado en el uso de “dígitos fríos”, que son los dígitos que han salido una sola vez o no han salido en un periodo de siete días. Él hizo la siguiente afirmación: “Ese [sistema] produce las combinaciones 5-8-9, 7-8-9, 6-8-9, 0-8-9 y 3-8-9. Estas cinco combinaciones tienen una excelente oportunidad de salir esta semana. Buena suerte para todos”. ¿Puede funcionar este sistema? ¿Por qué sí o por qué no?

Ejercicios de repaso del capítulo

Para los ejercicios 1 a 6 utilice los datos siguientes, obtenidos de un ensayo clínico de la prueba de sangre de Abbot, una prueba de sangre para embarazo (basada en información de “Specificity and Detection Limit of Ten Pregnancy Tests”, de Tiitinen y Stenman, *Scandinavian Journal of Clinical Laboratory Investigation*, vol. 53, Supp. 216).

	Resultado positivo de la prueba (indica embarazo)	Resultado negativo de la prueba (no indica embarazo)
Paciente está embarazada	80	5
Paciente no está embarazada	3	11

- Si 1 de los 99 sujetos se selecciona al azar, determine la probabilidad de obtener a una persona que esté embarazada.
- Si 1 de los 99 sujetos se selecciona al azar, determine la probabilidad de obtener un sujeto cuya prueba haya salido positiva.
- Si 1 de los 99 sujetos se selecciona al azar, determine la probabilidad de obtener una persona que esté embarazada o con prueba positiva.
- Si 1 de los 99 sujetos se selecciona al azar, determine la probabilidad de obtener una persona con prueba negativa o que esté embarazada.
- Si dos sujetos *diferentes* son elegidos aleatoriamente, determine la probabilidad de que ambos hayan dado positivo en la prueba.
- Si se seleccionan al azar dos personas *diferentes*, determine la probabilidad de que ambas estén embarazadas.
- La compañía Binary Computer fabrica circuitos para computadoras utilizados en los reproductores de DVD. Estos circuitos se producen con un rendimiento de 27%, lo que significa que 27% de ellos son buenos y el resto es defectuoso.
 - Si se selecciona un circuito al azar, determine la probabilidad de que éste *no* sea bueno.
 - Si se seleccionan dos circuitos al azar, determine la probabilidad de que ambos sean buenos.
 - Si se seleccionan al azar cinco circuitos, ¿cuál es el *número esperado* de circuitos buenos?
 - Si se eligen al azar cinco circuitos, determine la probabilidad de que todos sean buenos. Si hubiese obtenido los cinco circuitos buenos entre los cinco seleccionados, ¿continuaría creyendo que el rendimiento fue de 27%? ¿Por qué sí o por qué no?

Para los ejercicios 8 al 10 considere un evento como “raro” si su probabilidad es menor o igual a 0.05.

- Un instructor obsesionado con el sistema métrico insiste en que todas las preguntas de opción múltiple tengan 10 posibles respuestas diferentes, una de las cuales es correcta. ¿Cuál es la probabilidad de responder correctamente una pregunta si se hizo una elección al azar de la respuesta?
 - ¿Es “raro” responder correctamente una pregunta por adivinación?
- Una ruleta tiene 38 sectores: un sector es 0, otro es 00 y los demás están numerados del 1 al 36. Si apuesta todo su dinero al número 13 en un giro de la ruleta, ¿cuál es la probabilidad de que ganará?
 - ¿Es “raro” ganar cuando apuesta a un solo número en la ruleta?
- Un estudio de 400 vuelos seleccionados aleatoriamente de American Airlines mostró que 344 llegaron a tiempo. ¿Cuál es la probabilidad estimada de que un vuelo de American Airlines llegue demorado?
 - ¿Es “raro” para un vuelo de American Airlines llegar tarde?

Cuestionario del capítulo

1. Un examen de opción múltiple tiene respuestas de a, b, c, d, e y f, y sólo una respuesta es correcta. Si hace una elección al azar, ¿cuál es la probabilidad de que acierte a la respuesta correcta?
2. Tres preguntas de un examen de opción múltiple, cada una tiene respuestas de a, b, c, d, e y f una de las cuales es correcta en cada caso. Si hace una elección al azar para cada una de las tres respuestas, ¿cuál es la probabilidad de que las tres respuestas sean correctas?
3. Si $P(A) = 0.45$, determine $P(\bar{A})$.
4. Si se lanza 15 veces una moneda, determine el número esperado de caras.
5. En un ensayo clínico, algunos sujetos fueron tratados con Nicorette y a otros se les dio un placebo. Los resultados muestran una probabilidad de 0.279 de que una reacción adversa sea debida al azar. ¿El azar aparece como una explicación razonable para los resultados?

En los ejercicios 6 a 10 utilizan los datos siguientes de un ensayo clínico de Nicorette, una goma de mascar diseñada para ayudar a la gente a dejar de fumar (con base en información de Merrell Do Pharmaceuticals, Inc.).

	Nicorette	Placebo
Dolor en la boca o en la garganta	43	35
Sin dolor en la boca ni en la garganta	109	118

6. Si uno de los sujetos de la prueba se selecciona al azar, determine la probabilidad de obtener a alguien que se le dio un placebo.
7. Si uno de los sujetos de la prueba se selecciona al azar, determine la probabilidad de obtener a alguien que usó Nicorette y experimentó dolor en la boca o en la garganta.
8. Si uno de los sujetos de la prueba se selecciona al azar, determine la probabilidad de obtener a alguien que usó Nicorette o bien experimentó dolor en la boca o en la garganta.
9. Si se seleccionan dos sujetos diferentes al azar, determine la probabilidad de que ambos fueron del grupo placebo.
10. Si se eligen al azar tres sujetos diferentes, determine la probabilidad de que los tres son del grupo de tratamiento con Nicorette.

HABLEMOS DE CIENCIAS SOCIALES

¿Son justas las loterías?

Las loterías han sido parte de la forma de vida estadounidense. La mayoría de los estados tienen loterías legales, loterías multiestado, tales como Power-ball y Big Game. Las estadísticas nacionales muestran que el gasto per cápita (promedio por persona) en loterías es aproximadamente \$200 al año. Puesto que muchas personas no juegan en absoluto, esto significa que los jugadores activos tienden a gastar mucho más que \$200 anuales.

Las matemáticas de las posibilidades en la lotería incluyen conteo de diferentes combinaciones de números que son los ganadores. Aunque estos cálculos pueden ser muy complejos, la conclusión esencial siempre es la misma: la posibilidad de ganar un gran premio es infinitesimalmente pequeña. Los anuncios pueden hacer que la lotería suene como una buena inversión, pero el valor esperado asociado con una lotería siempre es negativo. En promedio, aquellos que juegan regularmente pueden esperar perder alrededor de la mitad de lo que gastan.

Los defensores de las loterías señalan varios aspectos positivos. Por ejemplo, las loterías producen miles de millones de dólares de ingreso que los estados utilizan para iniciativas en educación, recreación y medio ambiente. Estos ingresos permiten a los estados mantener tasas bajas de impuestos, más bajas de lo que serían de otro modo. Los defensores también señalan que la participación en la lotería es voluntaria y divertida para una muestra representativa de la sociedad. En realidad, una encuesta reciente de Gallup muestra que tres cuartos de los estadounidenses aprueban las loterías estatales (dos tercios aprueban los juegos legales en general).

Este panorama favorable es parte de la propaganda y las relaciones públicas de las loterías estatales. Por ejemplo, los funcionarios de la lotería de Colorado ofrecen estadísticas sobre la edad, ingresos y educación de los jugadores de lotería comparadas con las de la población general (figura 6.16). Dentro de escasos puntos porcentuales, la edad de los jugadores de lotería asemeja a la de la población como un todo. De manera análoga, el

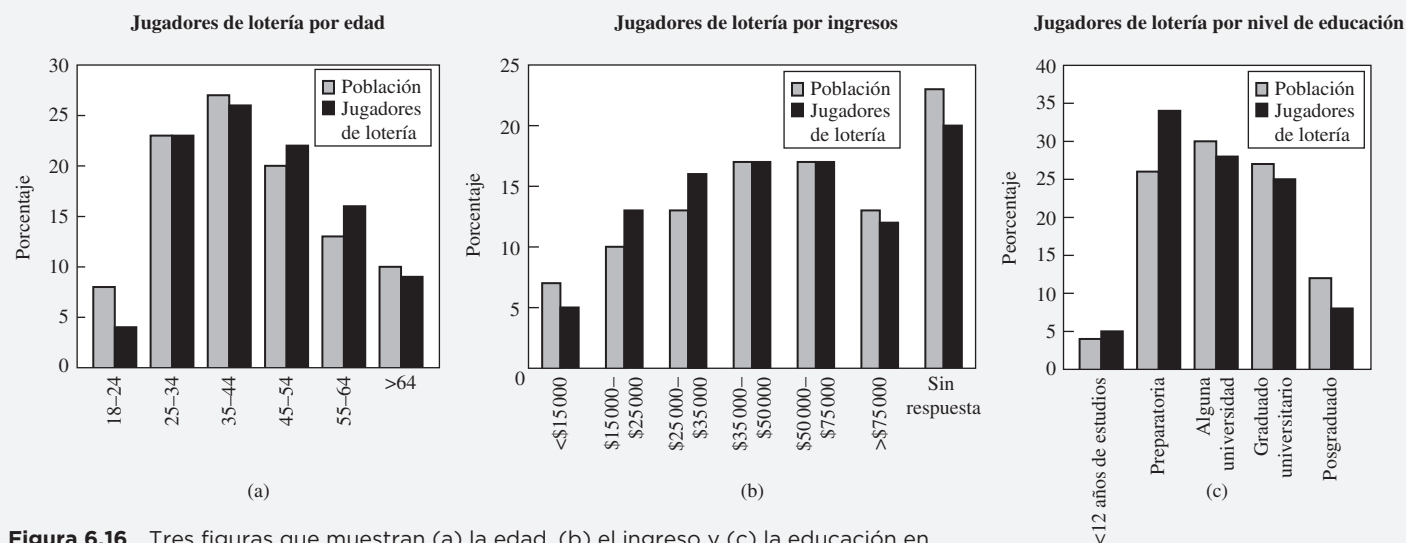


Figura 6.16 Tres figuras que muestran (a) la edad, (b) el ingreso y (c) la educación en jugadores de la lotería de Colorado en comparación con la población.

histograma para los ingresos de los jugadores de lotería da la impresión que éstos como un todo son ciudadanos típicos, con la excepción de las barras para ingresos de \$15 000-\$25 000 y \$25 000-\$35 000, lo cual muestra que los pobres tienden a gastar más de lo que esperaríamos por su proporción de la población.

No obstante, los beneficios aparentes de las loterías, los críticos han argumentado mucho que son sólo una forma desleal de impuestos. Un apoyo para esta opinión proviene de un reporte reciente de la Comisión Nacional de Estudio del Impacto de Juegos de Azar y un estudio del *New York Times* de las loterías en Nueva Jersey. Ambos estudios se centran en *el monto* de dinero gastado en loterías por los individuos.

El estudio del *New York Times* estuvo basado en datos de 48 875 personas quienes habían ganado al menos \$600 en los juegos de lotería de Nueva Jersey. (En un muestreo ingenioso, estos ganadores fueron tomados como una muestra aleatoria de todos los jugadores de lotería; después de todo, los ganadores de lotería son determinados aleatoriamente. Sin embargo, en realidad la muestra no es representativa de todos los jugadores de lotería ya que los ganadores tienden a comprar más del número promedio de billetes). Mediante la identificación de los códigos postales de los jugadores de lotería, los investigadores fueron capaces de determinar si los jugadores provenían de áreas con altos o bajos ingresos, alto o bajo promedio educativo y varias características demográficas. La conclusión aplastante del estudio del *New York Times* es que el gasto en lotería tiene un gran impacto *en términos relativos* en aquellos jugadores con más bajos ingresos y más bajo nivel educativo. Por ejemplo, lo siguiente estuvo entre los hallazgos específicos:

- La gente en las áreas del estado con ingresos más bajos gastan cinco veces más de sus ingresos en loterías que aquéllos en áreas del estado con ingresos más altos. El gasto en áreas de más bajos ingresos en un juego particular de lotería fue \$29 por cada \$10 000 de ingresos anuales, comparado con menos de \$5 por cada \$10 000 de ingresos anuales en las áreas de mayores ingresos.
- El número de expendios de lotería (donde los billetes de lotería pueden comprarse) es casi el doble por cada 10 000 personas en las áreas de bajos ingresos comparado con áreas de altos ingresos.
- La gente en áreas con los más bajos porcentajes de nivel educativo gastan más de cinco veces por cada \$10 000 de ingreso anuales que quienes se encuentran en áreas con porcentajes más elevados de educación.
- La publicidad y la promoción de loterías se concentra en áreas de bajos ingresos.

Algunos resultados del estudio de *New York Times* se resumieron en la figura 6.17. Sugiere que aunque Nueva Jersey tiene un sistema de impuestos progresivo (personas con mayores ingresos pagan un mayor porcentaje de sus ingresos en impuestos), el “impuesto de lotería” es regresivo. Además, el estudio encontró que las áreas que generan los mayores porcentajes de ingresos de la lotería no reciben una parte proporcional de fondos estatales.

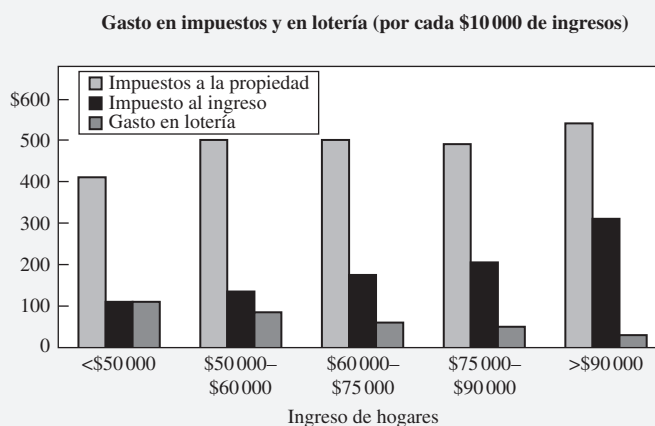


Figura 6.17 Gasto en impuestos y en lotería por ganadores de la lotería de Nueva Jersey (todas las cifras son por cada \$10 000 de ingresos).

Estudios similares revelan los mismos patrones en otros estados. Las conclusiones generales son ineludibles: aunque las loterías proporcionan muchos beneficios a los gobiernos estatales, el ingreso que producen proviene de manera desproporcionada de los más pobres e individuos con menos educación. En realidad, un reporte de la Comisión Nacional de Estudio del Impacto de Juegos de Azar concluyó que las loterías son “la forma más extendida de juegos en Estados Unidos” y que los gobiernos estatales han “penetrado de manera irresponsable con juegos de azar a la sociedad a una escala masiva... a través de medidas tales como publicidad incesante y colocación omnipresente de máquinas de lotería en tiendas del vecindario”.

PREGUNTAS PARA DISCUSIÓN

1. Estudie la figura 6.16. ¿Los jugadores de lotería parecen ser una muestra típica de la sociedad estadounidense con base en la edad? ¿Con base en los ingresos? ¿Con base en el nivel educativo? Explique. ¿Cómo afecta a estas conclusiones la categoría “sin respuesta”?
2. Algunos jugadores de lotería utilizan “sistemas” para la elección de números. Por ejemplo, consultan a “expertos” quienes les dicen (algunos por un pago) cuáles números son populares o deben de aparecer. En realidad, ¿tales sistemas pueden mejorar sus posibilidades de ganar la lotería? ¿Por qué sí o por qué no? ¿Considera que el uso de tales sistemas está relacionado con la formación educativa?
3. Considerando todos los factores presentados en esta sección y otros hechos que usted pueda encontrar, ¿considera que las loterías son justas para los pobres y para la gente sin educación? ¿Deben seguir siendo legales? ¿Deben restringirse de alguna manera?
4. Encuentre y estudie un anuncio de lotería y determine si de alguna manera es engañoso.
5. Una cita anónima que circuló en internet decía: “Las loterías son un impuesto para gente que no sabe matemáticas”. Comente el significado y la precisión de esta cita.

LECTURAS RECOMENDADAS

Pulley, Brett, “Living Off the Daily Dream of Winning a Lottery Prize”, *New York Times*, 22 de mayo de 1999.

Safire, William, “Lotteries Are Losers”, *New York Times*, 21 de junio de 1999.

Schwartz, Nelson D., y Nixon, Ron, “Some States Consider Leasing Their Lotteries”, *New York Times*, 14 de octubre de 2007.

Stodghill, Ron y Nixon, Ron, “For Schools, No Lucky Numbers”, *New York Times*, 7 de octubre de 2007.

Walsh, James, *True Odds*, Merrin Publishing, 1996.



¿El ADN es una huella confiable?

La huella ADN (también llamada perfil ADN o identificación ADN) se ha convertido en una herramienta importante en la aplicación de la ley. Es utilizada en casos criminales, en casos de paternidad e incluso en la identificación de restos humanos. (La huella ADN fue la principal forma en que restos de víctimas fueron identificados después de los ataques terroristas del 11 de septiembre en el World Trade Center en 2001).

El fundamento científico para la identificación del ADN ha sido conocido por varias décadas. Sin embargo, estas ideas fueron usadas por primera vez en la sala de juicios sólo hasta 1986. El caso involucró a un joven de 17 años acusado de violar y asesinar a dos estudiantes en Narborough, en la región central de Inglaterra. Durante su interrogatorio al sospechoso se le hizo una prueba de sangre, que fue enviada al laboratorio de un connotado genetista, Alec Jeffreys, en la cercana Universidad de Leicester. Usando métodos que él ya había desarrollado para pruebas de paternidad, Jeffreys comparó el ADN del sospechoso con el que se encontró en muestras de las víctimas. Las pruebas mostraron que las violaciones fueron cometidas por la misma persona, pero no por el sospechoso en custodia. El año siguiente, después de que más de 4500 muestras de sangre fueron recolectadas, los investigadores hicieron una identificación positiva del asesino usando los métodos de Jeffreys.

La noticia del caso británico y de los métodos de Jeffreys se extendieron velozmente. Las técnicas fueron rápidamente probadas, comercializadas y promovidas. Como era de esperar, el uso de identificación del ADN también encontró de inmediato oposición y controversia.

Para explorar los papeles esenciales que desempeñan la probabilidad y la estadística en la identificación del ADN, considere una analogía sencilla de testigo presencial. Suponga que está buscando a una persona quien le ayudó durante un momento de necesidad y sólo recuerda tres cosas acerca de esta persona:

- La persona era mujer.
- Ella tenía ojos verdes.
- Ella era pelirroja de cabello largo.

Si encuentra a alguien que coincide con este perfil, ¿podría concluir que esta persona es la mujer que le ayudó? Para responder, necesita información que le diga las probabilidades de que al seleccionar de manera aleatoria individuos en la población tengan estas características. La probabilidad de que una persona sea mujer es alrededor de $1/2$. Digamos que la probabilidad de ojos verdes es alrededor de 0.06 (6% de la población tiene ojos verdes) y la probabilidad de cabello rojo largo es 0.0075. Si suponemos que estas características ocurren de manera independiente una de las otras, entonces la probabilidad que una persona seleccionada tenga las tres características es

$$0.5 \times 0.06 \times 0.0075 = 0.000225$$

o alrededor de 2 en 10000. Esto puede parecer relativamente bajo, pero es probable que no sea lo suficientemente bajo para sacar una conclusión definitiva. Por ejemplo, un perfil satisfecho por sólo 2 en 10000 personas seguirá siendo satisfecho por alrededor de 200 personas en una ciudad con una población en 1 millón. Claro, si quiere estar seguro que ha encontrado a la persona correcta, necesitará información adicional para el perfil.

La identificación del ADN tiene como base la misma idea, pero está diseñada de modo que la probabilidad que un perfil coincida sea mucho menor. El ADN de cada individuo es

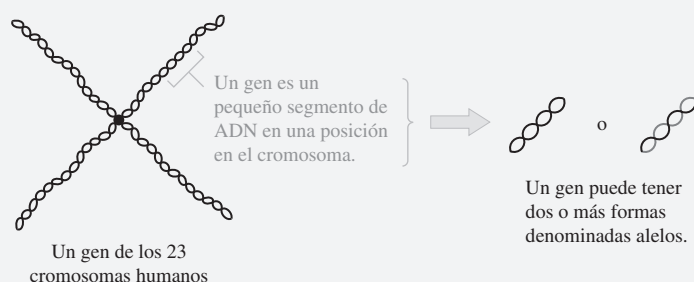


Figura 6.18 Diagrama que muestra la relación entre cromosomas, alelos, genes y posiciones.

único y es el mismo en todo el individuo. Un solo rasgo físico está determinado por una pequeña pieza de ADN llamada *gen* en una *posición* (ubicación) en un *cromosoma*; los humanos tienen 23 cromosomas y más de 30 000 genes (figura 6.18). Un gen puede tomar dos o más (con frecuencia cientos) de formas diferentes denominadas *alelos*. Alelos diferentes dan lugar a variaciones de un rasgo (por ejemplo, diferente color de cabello o diferente tipo de sangre). No sólo pueden aparecer diferentes alelos en una posición, sino la pieza de ADN puede tener longitudes diferentes en diferentes personas (denominada *número variable de secuencias repetidas* o VNTR, por sus siglas en inglés). La evidencia genética que es recolectada y analizada en el laboratorio consiste de las longitudes de los alelos o tipos de alelos en cinco a ocho posiciones diferentes.

La evidencia genética reunida (de muestras de sangre, tejido, cabellos, semen, o incluso saliva en una estampilla postal) y la analizada son directas, al menos en teoría. Sin embargo, el proceso está sujeto tanto a controversia como a fuentes de error. Suponga, en nuestra analogía, que se encuentra a una persona con el cabello “rojo cobrizo” en lugar de rojo. Puesto que muchas características son variables continuas (no discretas), ¿debe descartar a esta persona o suponer que el rojo cobrizo es suficientemente cercano? Por esta razón, el tema de *clasificación* se vuelve extremadamente importante (vea la sección 3.1). Podría elegir incluir a todas las personas con cabello con cualquier tonalidad de rojo, o podría elegir una clase más cerrada —digamos, sólo cabello rojo vivo— que daría una prueba más discriminatória.

El mismo tema surge en pruebas genéticas. Cuando se miden los tipos o las longitudes de los alelos en el laboratorio, hay suficiente variabilidad o error en las mediciones, ya que estas variables son continuas. Los anchos de clase deben elegirse, y la elección es la fuente del debate. Las clases con un ancho pequeño dan pruebas más refinadas, excluyen a más sospechosos, y en última instancia proporcionan evidencia más sólida frente a un acusado.

Otra fuente de controversia científica es la suposición de que las características en un perfil genético son independientes. Testigos expertos (genetistas y biólogos moleculares) han testificado en largos debates acerca de este punto sin coincidir. Parece que se acuerda en que la suposición de independencia no introduce errores significativos en vista de otras fuentes de error. Pero es importante ver cómo una pregunta científica difícil afecta a la matemática involucrada: si las características son independientes, entonces la regla de la multiplicación para eventos independientes está justificada; si no, un tipo diferente de probabilidad necesita usarse.

Un tercer punto de controversia concierne a la elección de una población de referencia. Algunos argumentan que si se utiliza la información genética de una población completa, podría no representar de manera justa a las subpoblaciones étnicas. Por ejemplo, un alelo particular podría tener una frecuencia muy diferente en la minoría étnica italiana que en toda la población de Estados Unidos. Tal discrepancia cambiaría el resultado del cálculo de la probabilidad y podría fortalecer o debilitar el caso de un acusado.

La figura 6.19 muestra un ejemplo de información de la población de referencia usada en pruebas genéticas. El eje horizontal muestra 30 clases para las diferentes medidas de los alelos y el eje vertical muestra la frecuencia para cada clase. De igual interés son las curvas de frecuencia para cuatro subpoblaciones asiáticas. La variación significativa entre las curvas es la evidencia de usar bases de datos especializadas para subpoblaciones. Sin embargo, hasta que tales bases de datos estén completas y exista evidencia de apoyo en un caso que utilice una base de datos especializada, la opción aún es utilizar la población completa como referencia.

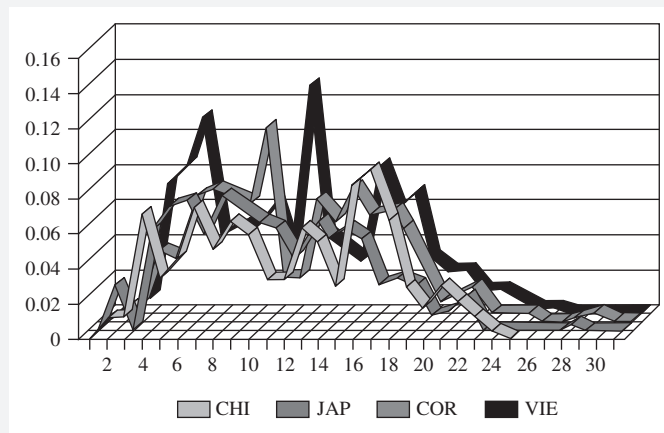


Figura 6.19 Datos de frecuencia por clases para un alelo, que muestra la variabilidad entre cuatro grupos étnicos. *Fuente:* Kathryn Roeder. “DNA Fingerprinting: A Review of the Controversy”, *Statistical Science*, volumen 9, número 2. pp. 222-247

La tecnología del ADN ha creado nuevos caminos que entrelazan a científicos y especialistas en leyes que están cambiando nuestra sociedad. Como era de esperar, los métodos y el pensamiento estadísticos son una parte esencial del cambio.

PREGUNTAS PARA DISCUSIÓN

1. El resultado de una prueba de ADN es considerada una prueba física (en contraposición con la circunstancial). Aun es mucho más sofisticada y difícil de entender que una pieza típica de evidencia física, tal como una pistola o parte de una ropa. Por tanto, algunas personas han argumentado que la evidencia del ADN no debe usarse en casos criminales en los que los miembros del jurado no comprendan completamente cómo la evidencia fue tomada y analizada. ¿Qué piensa de este argumento? Defienda su opinión.
2. Suponga que un alelo tiene una frecuencia *mayor* en una subpoblación que en la población completa. Explique cómo esto cambiaría la evidencia a favor o en contra de un acusado. Suponga que un alelo tiene una frecuencia *menor* en una subpoblación que en la población completa. Explique cómo esto cambiaría la evidencia a favor o en contra de un acusado.
3. La evidencia de pruebas de sangre pueden identificar a sospechosos con una probabilidad de casi 1 en 200. La evidencia de pruebas de ADN con frecuencia proporciona probabilidades que se asegura son del orden de 1 en 10 millones. Si usted fuese miembro de un jurado, ¿aceptaría tal probabilidad como una identificación positiva de un sospechoso?
4. El Proyecto Inocencia utiliza el ADN para tratar de aclarar la situación de sospechosos injustamente convictos de crímenes. ¿Es más sencillo establecer la inocencia que la culpabilidad por medio de pruebas de ADN? Explique.
5. Analice otras maneras en que pruebas de ADN pueden ser útiles, tales como en temas de resolución de paternidad. Globalmente, ¿cuánto considera que la evidencia del ADN es probable que afecte a nuestra sociedad en el futuro?

LECTURAS RECOMENDADAS

La tecnología de huellas de ADN de manera constante se está refinando y actualizando. Puede encontrar una gran cantidad de información buscando en la web.

Aprenda más acerca del Proyecto Inocencia (Project Innocence) en <http://www.innocence-project.org>.



*La persona que sabe “cómo”,
siempre tendrá un trabajo.
La persona que sabe “por qué”,
siempre será su jefe.*

—Diane Ravitch

Correlación y causalidad

¿FUMAR CAUSA CÁNCER DE PULMÓN? ¿BAJO DESEMPLEO

lleva a la inflación? ¿El uso humano de combustibles fósiles provoca el calentamiento global? Un objetivo principal de muchas investigaciones estadísticas es buscar las relaciones entre las diferentes variables, de modo que los investigadores pueden determinar si un factor *causa* al otro. Una vez descubierta una relación, podemos tratar de determinar si existe una causa subyacente. En este capítulo estudiaremos las correlaciones y exploraremos su importancia para la tarea más difícil de buscar la causalidad.

OBJETIVOS DE APRENDIZAJE

7.1 Búsqueda de correlación

Ser capaz de definir correlación, reconocer correlaciones positiva y negativa en diagramas de dispersión, y comprender el coeficiente de correlación como una medida de la intensidad de una correlación.

7.2 Interpretación de correlaciones

Tomar en cuenta las precauciones debidas en la interpretación de la correlación, en especial los efectos de los datos atípicos, la agrupación de datos y, el hecho crucial, que la correlación no necesariamente implica causalidad.

7.3 Rectas de mejor ajuste y pronóstico

Familiarizarse con el concepto de una recta de mejor ajuste para una correlación, reconociendo cuándo tales rectas tienen un valor predictivo y cuándo no, entender cómo el cuadrado del coeficiente de correlación está relacionado con la calidad del ajuste, y entender cualitativamente el uso de regresión múltiple.

7.4 La búsqueda de la causalidad

Entender la dificultad de establecer causalidad a partir de la correlación, e investigar directrices que puedan usarse para ayuda a establecer confianza en la causalidad.

7.1 Búsqueda de correlación

Fumar es una de las causas principales de la estadística.

—Fletcher Knebel

• Qué significa cuando decimos que fumar *causa* cáncer de pulmón? Seguro *no* significa que usted tendrá cáncer de pulmón si fuma un solo cigarro. E incluso no significa que definitivamente tendrá cáncer de pulmón si fuma mucho durante muchos años, ya que algunos fumadores empedernidos no adquieren cáncer de pulmón. En cambio, es un enunciado *estadístico* que significa que es *mucho más probable* que usted adquiera cáncer de pulmón si fuma que si no fuma.

Los investigadores, ¿cómo supieron que fumar causa cáncer de pulmón? El proceso inició con observaciones informales, cuando los doctores observaron que una sorprendentemente alta proporción de sus pacientes con cáncer de pulmón eran fumadores. Estas observaciones condujeron a estudios cuidadosos en que los investigadores compararon las tasas de cáncer de pulmón entre fumadores y no fumadores. Dichos estudios mostraron con claridad que era más probable que fumadores empedernidos adquiriesen cáncer de pulmón. En términos más formales, decimos que existe una **correlación** entre las variables *cantidad de fumar* y *posibilidad de cáncer de pulmón*. Una correlación es un tipo especial de relación entre variables, en la que un aumento o disminución de una lleva consigo un correspondiente aumento o disminución en la otra.

Definición

Una **correlación** existe entre dos variables cuando valores mayores de una variable van de manera consistente con valores mayores de otra variable o cuando valores grandes de una variable corresponden de manera consistente a valores menores de otra variable.

A continuación, algunos ejemplos de correlaciones:

- Existe una correlación entre las variables *estatura* y *peso* de una persona; esto es, personas más altas tienden a pesar más que personas más bajas.
- Existe una correlación entre las variables *demanda de manzanas* y *precio de las manzanas*; esto es, la demanda tiende a disminuir cuando el precio aumenta.
- Existe una correlación entre *tiempo de práctica* y *habilidad* entre ejecutantes de piano; esto es, aquellos que practican más tienden a tener más habilidad.

Es importante darse cuenta que establecer una correlación entre dos variables *no* significa que un cambio en una variable *cause* un cambio en la otra. La correlación entre fumar y cáncer de pulmón no está probando que fumar cause cáncer de pulmón. Por ejemplo, podríamos imaginar que algún gen predispone a una persona tanto a fumar como a tener cáncer de pulmón. Sin embargo, la identificación de la correlación fue el primer paso crucial en el aprendizaje de que fumar causa cáncer de pulmón. Posteriormente, en este capítulo analizaremos la difícil tarea de establecer causalidad. Por ahora, nos concentramos en cómo buscamos, identificamos e interpretamos correlaciones.

UN MOMENTO DE REFLEXIÓN

Suponga que en realidad fuese un gen el que hace a la gente propensa a fumar y a tener cáncer de pulmón. Explique por qué aún encontraríamos una fuerte correlación entre fumar y cáncer de pulmón en ese caso, pero no podríamos decir que fumar causa cáncer de pulmón.

A propósito...

Fumar está ligado a muchas enfermedades graves además de cáncer de pulmón, incluyendo cardiopatías y enfisema. Fumar también está ligado con menos condiciones mortales de salud tales como arrugas prematuras en la piel e impotencia sexual.

Diagramas de dispersión

La tabla 7.1 lista datos para una muestra de diamantes en una joyería, sus precios y varias medidas comunes que ayudan a determinar su valor. Puesto que los anuncios para diamantes con frecuencia sólo indican sus pesos (quilates), podríamos sospechar una correlación entre los pesos y los precios. Podemos buscar tal correlación haciendo un **diagrama de dispersión** (o *gráfica de dispersión*) que muestre la relación entre las variables *peso* y *precio*.

Tabla 7.1 Precios y características de una muestra de 23 diamantes de un distribuidor de gemas

Diamante	Precio	Peso (quilates)	Profundidad	Mesa	Color	Claridad
1	\$6 958	1.00	60.5	65	3	4
2	\$5 885	1.00	59.2	65	5	4
3	\$6 333	1.01	62.3	55	4	4
4	\$4 299	1.01	64.4	62	5	5
5	\$9 589	1.02	63.9	58	2	3
6	\$6 921	1.04	60.0	61	4	4
7	\$4 426	1.04	62.0	62	5	5
8	\$6 885	1.07	63.6	61	4	3
9	\$5 826	1.07	61.6	62	5	5
10	\$3 670	1.11	60.4	60	9	4
11	\$7 176	1.12	60.2	65	2	3
12	\$7 497	1.16	59.5	60	5	3
13	\$5 170	1.20	62.6	61	6	4
14	\$5 547	1.23	59.2	65	7	4
15	\$7 521	1.29	59.6	59	6	2
16	\$7 260	1.50	61.1	65	6	4
17	\$8 139	1.51	63.0	60	6	4
18	\$12 196	1.67	58.7	64	3	5
19	\$14 998	1.72	58.5	61	4	3
20	\$9 736	1.76	57.9	62	8	2
21	\$9 859	1.80	59.6	63	5	5
22	\$12 398	1.88	62.9	62	6	2
23	\$11 008	2.03	62.0	63	8	3

Notas: el peso se mide en quilates (1 quilate = 0.2 gramos). La profundidad se define como 100 veces la razón de la altura al diámetro. La mesa es el tamaño de la superficie plana superior. (La profundidad y la mesa determinan el “corte”). El color y la claridad, cada una, son medidas en escalas estándar, donde 1 es mejor. Para color, 1 = sin color y números en aumento indican más amarillo. Para claridad 1 = sin fallas y 6 indica que los defectos pueden verse a simple vista.

A propósito...

La palabra *quilate* usada para describir oro no tiene el mismo significado que el término *quilate* para diamantes y otras gemas. Un quilate (para gemas) es una medida de peso igual a 0.2 gramos. Los quilates (para oro) son una medida de la pureza del oro: oro de 24 quilates es 100% oro puro; oro de 18 quilates es 75% oro puro (y 25% de otros metales); oro de 12 quilates es 50% oro puro (y 50% de otros metales), y así sucesivamente.



Definición

Un **diagrama de dispersión** (o *gráfica de dispersión*) es una gráfica en la que cada punto representa los valores de dos variables.

El procedimiento siguiente describe cómo construir el diagrama de dispersión de la figura 7.1.

1. Asignamos una variable a cada eje y rotulamos el eje con los valores que queden de manera espaciada para que quepan todos los datos. En ocasiones la selección de los ejes es arbitraria, pero si sospecha que una variable depende de la otra, entonces trace la *variable explicativa* en el eje horizontal y la *variable de respuesta* en el eje vertical. En este caso esperamos que el precio del diamante dependa, al menos en parte, de su peso; por tanto, decimos que el *peso* es la variable explicativa (ya que ayuda a *explicar* el precio) y el *precio* es la variable de respuesta (ya que *responde* a cambios en la variable explicativa). Elegimos el rango de 0 a 2.5 quilates para el eje del *peso* y \$0 a \$16 000 para el eje del *precio*.
2. Para cada diamante en la tabla 7.1 trazamos un *solo punto* en la posición horizontal que corresponde a su peso y en la posición vertical que corresponde a su precio. Por ejemplo, el punto para el diamante 10 va en la posición de 1.11 quilates en el eje horizontal y \$3 670 en el eje vertical. Las líneas discontinuas en la figura 7.1 muestran cómo localizar este punto.

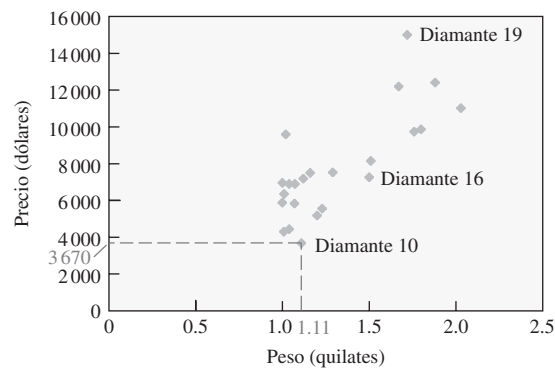


Figura 7.1 Diagrama de dispersión que muestra la relación entre las variables *precio* y *peso* para los diamantes de la tabla 7.1. Las líneas discontinuas muestran cómo encontramos la posición del punto para el diamante 10.

3. (Opcional) Podemos rotular algunos (o todos) los puntos de datos, como se hizo para los diamantes 10, 16 y 19 en la figura 7.1.

Los diagramas de dispersión obtienen su nombre de la manera en que los puntos están esparcidos (o dispersos) y pueden revelar una relación entre las variables. En la figura 7.1 vemos una tendencia general ascendente que indica que los diamantes con mayor peso tienden a ser más caros. La correlación no es perfecta. Por ejemplo, el diamante más pesado no es el más caro. Pero la tendencia global parece ser muy clara.

UN MOMENTO DE REFLEXIÓN

Identifique los puntos en la figura 7.1 que representan a los diamantes 3, 7 y 23.

EJEMPLO 1 Color y precio

Con los datos de la tabla 7.1 cree un diagrama de dispersión para buscar una correlación entre el *color* de un diamante y el *precio*. Comente sobre la correlación.

Solución Esperamos que el precio dependa del color, de modo que graficamos la variable explicativa *color* en el eje horizontal y la variable de respuesta *precio* en el eje vertical en la figura 7.2. (Debe verificar algunos puntos contra los datos en la tabla 7.1). Los puntos parecen mucho más dispersos que en la figura 7.1. Sin embargo, puede notar una débil tendencia hacia abajo de la parte superior izquierda hacia la inferior derecha. Esta tendencia representa una correlación débil en que los diamantes con más color amarillo (números altos para la variable color) son menos caros. Esto es consistente con lo que esperaríamos, ya que la ausencia de

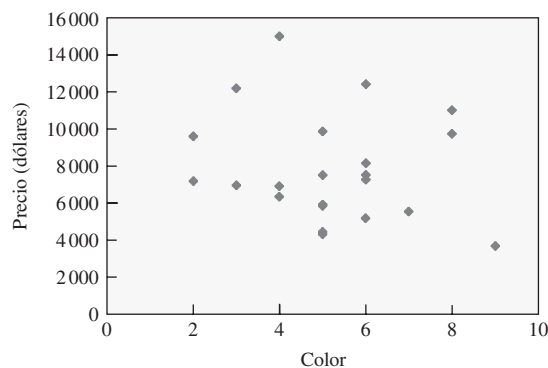


Figura 7.2 Diagrama de dispersión para la información de color y precio en la tabla 7.1.

colorido del diamante hace que parezcan más brillantes y en general son considerados más deseables.

UN MOMENTO DE REFLEXIÓN

Gracias a un gran bono en su trabajo, tiene un presupuesto de \$6000 para un anillo de diamante. Por ese precio un proveedor le ofrece las dos opciones siguientes: un diamante que pesa 1.20 quilates y tiene color = 4; el otro pesa 1.18 quilates y tiene color = 3. Si las demás características de los diamantes son iguales, ¿cuál elegiría? ¿Por qué?

Tipos de correlación

Hemos visto dos ejemplos de correlación. La figura 7.1 muestra una correlación bastante fuerte entre peso y precio, mientras que la figura 7.2 muestra una correlación débil entre color y precio. Ahora estamos preparados para generalizar los tipos de correlación. La figura 7.3 muestra

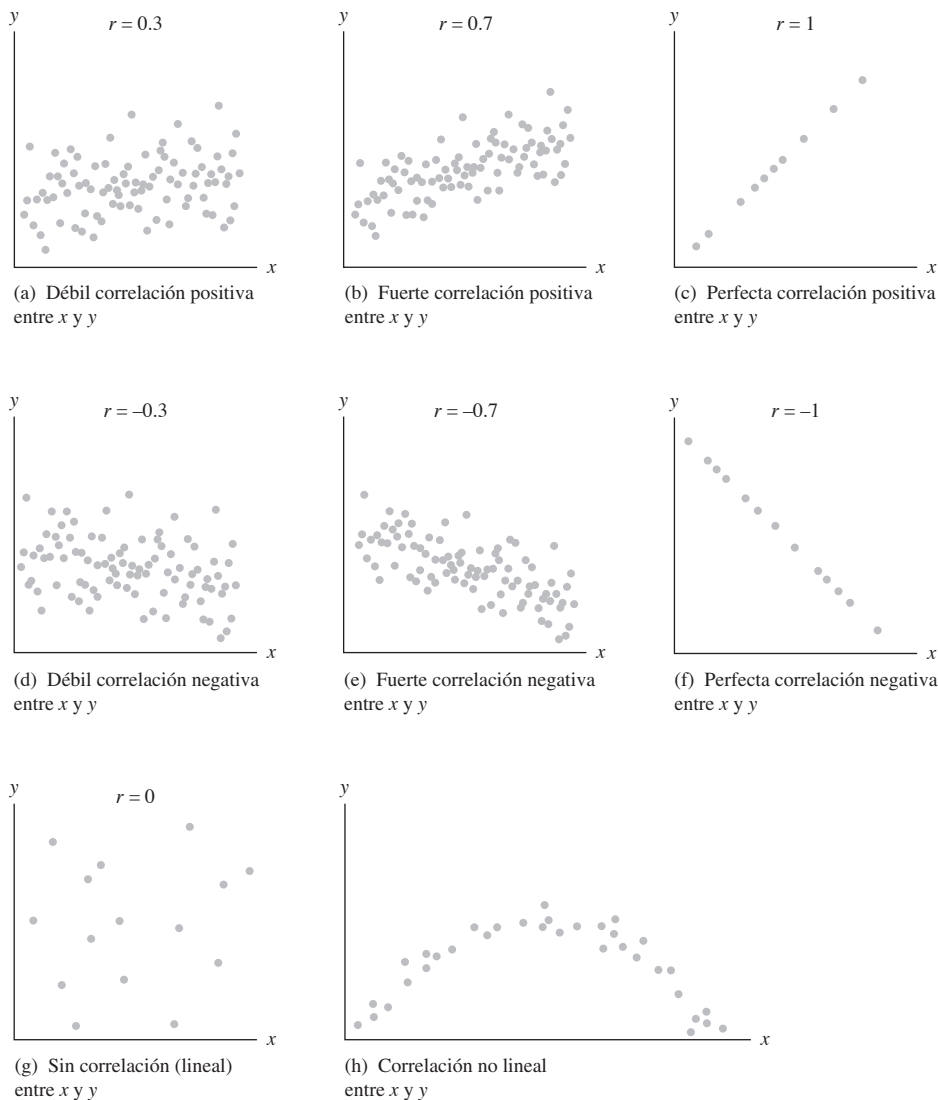


Figura 7.3 Tipos de correlación vistos en diagramas de dispersión.

ocho diagramas de dispersión para las variables x y y . Observe las siguientes características clave de estos diagramas:

- Las partes de a a c de la figura 7.3 muestran **correlaciones positivas**, en las cuales los valores de y tienden a aumentar con valores crecientes de x . La correlación se hace más fuerte conforme pasamos de a a c. De hecho, c muestra una perfecta correlación positiva, en la que todos los puntos caen a lo largo de una línea recta.
- Las partes d a f de la figura 7.3 muestran **correlaciones negativas**, en las cuales los valores de y tienden a disminuir con valores crecientes de x . La correlación se hace más fuerte conforme pasamos de d a f. De hecho, f muestra una perfecta correlación negativa, en la que todos los puntos caen a lo largo de una línea recta.
- La parte g de la figura 7.3 muestra una **no correlación** entre x y y . En otras palabras, los valores de x no parece que están ligados a los valores de y de alguna manera.
- La parte h de la figura 7.3 muestra una **relación no lineal**, en la cual x y y parecen estar relacionados, pero la relación no corresponde a una línea recta. (*Lineal* significa a lo largo de una línea recta y *no lineal* significa *no* a lo largo de una línea recta).

NOTA TÉCNICA

En este texto utilizamos el término *correlación* sólo para relaciones *lineales*. Algunos estadísticos se refieren a relaciones no lineales como “correlaciones no lineales”. Existen técnicas para el trabajo con relaciones no lineales que son similares a las que se describen en este libro para las relaciones lineales.

Tipos de correlación

Correlación positiva: ambas variables tienden a aumentar (o disminuir) al mismo tiempo.

Correlación negativa: las dos variables tienden a cambiar en direcciones opuestas, con una aumentando mientras la otra decrece.

No correlación: no existe una aparente relación (lineal) entre las dos variables.

Relación no lineal: las dos variables están relacionadas, pero la relación que resulta en un diagrama de dispersión no sigue un patrón de una línea recta.

EJEMPLO 2 Esperanza de vida y mortalidad infantil

La figura 7.4 muestra un diagrama de dispersión para las variables *esperanza de vida* y *mortalidad infantil* en 16 países. ¿Qué tipo de correlación muestra? ¿Esta correlación tiene sentido? ¿Implica causalidad? Explique.

Solución El diagrama muestra una moderada correlación negativa en la que los países con *baja* mortalidad infantil tienden a tener *alta* esperanza de vida. Es una correlación *negativa* ya que las dos variables cambian en direcciones opuestas. La correlación tiene sentido ya que

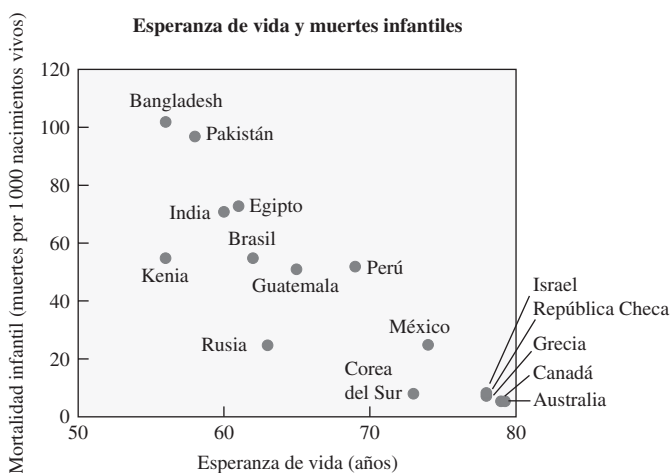


Figura 7.4 Diagrama de dispersión para la información de esperanza de vida y mortalidad infantil.

esperaríamos que países con mejores cuidados de salud tendrían baja mortalidad infantil y alta esperanza de vida. Sin embargo, *no* implica causalidad entre mortalidad infantil y esperanza de vida. No esperaríamos que un esfuerzo conjunto para reducir la mortalidad infantil aumentara la esperanza de vida de manera significativa, a menos que fuese parte de un esfuerzo global para mejorar el cuidado de la salud. (Al reducir la mortalidad infantil aumenta *ligeramente* la esperanza de vida, ya que al tener menos muertes infantiles se tiende a elevar la edad media de la muerte de la población).

Medidas de fuerza de una correlación

Para la mayoría de los propósitos es suficiente establecer si una correlación es fuerte, débil o no existe. Sin embargo, en ocasiones es útil describir la fuerza de una correlación en términos más precisos. Los estadísticos miden la fuerza de una correlación con el **coeficiente de correlación**, representado por la letra r .

Los coeficientes de correlación son fáciles de interpretar, aunque son tediosos de calcular, a menos que utilice una calculadora o una computadora. Vea nuevamente la figura 7.3 y observe que muestra el valor del coeficiente de correlación para cada uno de los diagramas de dispersión. El coeficiente de correlación siempre está entre -1 y 1 . Cuando los puntos en un diagrama de dispersión están cercanos a una recta ascendente, el coeficiente de correlación es positivo y cercano a 1 . De manera análoga, los puntos que están cercanos a una recta descendente tienen un coeficiente de correlación negativo con un valor cercano a -1 . Los puntos que no se ajustan a ningún tipo de patrón de recta o que están cercanos a una recta *horizontal* (indicando que los valores de y no dependen de los valores de x) dan lugar a un coeficiente de correlación cercano a 0 .

Propiedades del coeficiente de correlación, r

- El coeficiente de correlación, r , es una medida de la fuerza de una correlación; su valor puede variar sólo de -1 a 1 .
- Si no existe correlación, los puntos no siguen ningún patrón lineal ascendente o descendente, y el valor de r es cercano a cero.
- Si existe una correlación positiva, el coeficiente de correlación es positivo ($0 < r \leq 1$). Ambas variables aumentan juntas. Una perfecta correlación positiva (en la que todos los puntos en un diagrama de dispersión están en una recta ascendente) tiene un coeficiente de correlación de $r = 1$. Los valores de r cercanos a 1 significan una fuerte correlación positiva y valores positivos cercanos a 0 significan una débil correlación positiva.
- Si existe una correlación negativa, el coeficiente de correlación es negativo ($-1 \leq r < 0$): cuando una variable aumenta la otra disminuye. Una perfecta correlación negativa (en la que todos los puntos en un diagrama de dispersión están en una recta descendente) tiene un coeficiente de correlación de $r = -1$. Los valores de r cercanos a -1 significan una fuerte correlación negativa y valores negativos cercanos a 0 significan una débil correlación negativa.

NOTA TÉCNICA

Para los métodos de esta sección, existe el requisito de que las dos variables den como resultado datos que tengan una “distribución normal bivariada”. Esto básicamente significa que para cualquier valor fijo de una variable, los valores correspondientes de la otra variable tienen distribución normal. Este requerimiento por lo regular es muy difícil de verificar, por lo que la comprobación se reduce a verificar que ambas variables den como resultado datos que se distribuyan normalmente.

EJEMPLO 3 Tamaño de granjas en Estados Unidos

La figura 7.5 muestra un diagrama de dispersión para las variables *número de granjas* y *tamaño medio de la granja* en Estados Unidos. Cada punto representa datos de un solo año entre 1950 y 2000; en este diagrama, los primeros años por lo general están a la derecha y los últimos años están a la izquierda. Estime el coeficiente de correlación comparando este diagrama con los de la figura 7.3 y analice las razones subyacentes para la correlación.

Solución El diagrama de dispersión muestra una fuerte correlación negativa que se parece mucho al diagrama de dispersión en la figura 7.3f, lo que sugiere un coeficiente de correlación alrededor de $r = -0.9$. La correlación muestra que cuando el número de granjas disminuye, el

Apropósito...

En 1900 más del 40% de la población de Estados Unidos trabajaba en granjas; para 2000, menos de 2% de la población trabajaba en granjas.

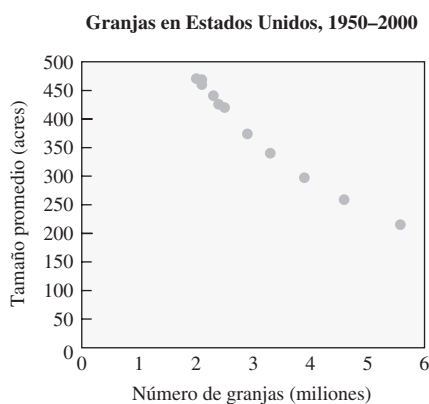


Figura 7.5 Diagrama de dispersión para los datos de tamaño de granjas. *Fuente:* Departamento de Agricultura, Estados Unidos.

tamaño de las granjas que quedan aumenta. Esta tendencia refleja un cambio básico en la naturaleza de la agricultura: antes de 1950 la mayoría de las granjas eran pequeñas granjas familiares. Con el paso del tiempo, estas pequeñas granjas han sido reemplazadas por grandes granjas pertenecientes a compañías de la agroindustria.

EJEMPLO 4 Precisión de los pronósticos del clima

Los diagramas de dispersión en la figura 7.6 muestran dos semanas de información que comparan las temperaturas máximas reales del día con el pronóstico del mismo día (parte a) y el pronóstico a tres días (parte b). Estime el coeficiente de correlación para cada conjunto de datos y analice lo que estos coeficientes implican con respecto a los pronósticos del clima.

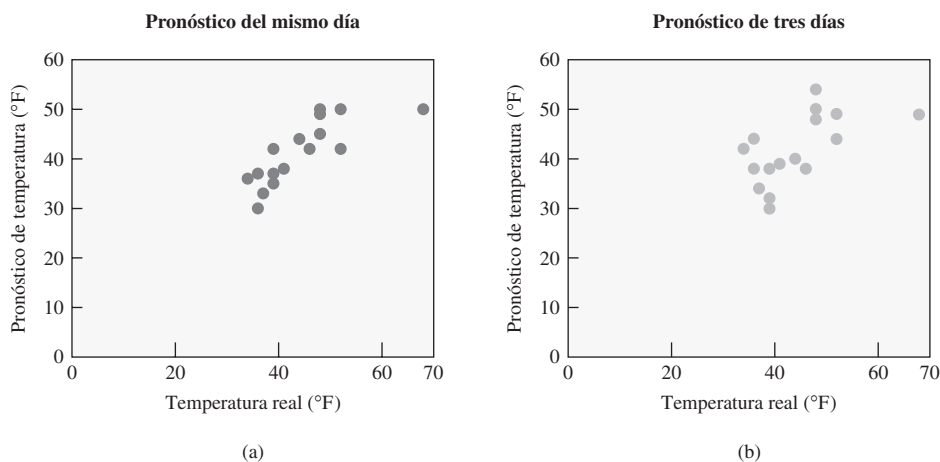


Figura 7.6 Comparación de temperaturas máximas reales con pronósticos de (a) el mismo día y (b) tres días.

Solución Si cada pronóstico fuese perfecto, cada temperatura real sería igual a la temperatura correspondiente pronosticada. Esto daría como resultado que todos los puntos estuviesen en una línea recta y un coeficiente de correlación de $r = 1$. En la figura 7.6a, en la que los pronósticos fueron realizados al inicio del mismo día, los puntos están muy cercanos a una recta, lo que significa que pronósticos del mismo día están muy relacionados con las temperaturas reales. Al comparar este diagrama de dispersión con los diagramas en la figura 7.3, podemos estimar, de manera razonable, este coeficiente de correlación alrededor de $r = 0.8$. La correlación es más débil en la figura 7.6b, indicando que los pronósticos hechos con tres días de anticipación no son tan cercanos a las temperaturas reales como los pronósticos del mismo día.

Este coeficiente de correlación es alrededor de 0.6. Estos resultados son comprensibles, ya que esperamos que pronósticos a largo plazo sean menos precisos.

EJEMPLO 5 Éxito de película

La tabla 7.2 muestra el costo de producción y recaudación bruta para las 15 películas de fantasía y de ficción con los más altos presupuestos de todos los tiempos (hasta 2006). Construya un diagrama de dispersión para la relación entre costo de producción y recaudación bruta. Estime el coeficiente de correlación y analice su significado. (Recaudación bruta es el monto total de dinero recolectado en la venta de boletos en las taquillas del cine).

Tabla 7.2 Películas de fantasía y ciencia ficción con presupuestos más grandes		
Película	Costo de producción (millones de dólares)	Recaudación bruta (millones de dólares)
King Kong (2005)	207	218
Hombre Araña 2 (2004)	200	373
Crónicas de Narnia (2005)	180	292
Waterworld (1995)	175	88
Van Helsing (2004)	170	120
El Expreso Polar (2004)	170	172
Terminator 3 (2003)	170	150
Poseidón (2006)	160	52
Batman: el inicio (2005)	150	205
Harry Potter/El cáliz de fuego (2005)	150	290
Armagedón (1998)	140	201
Hombres de negro 2 (2002)	140	190
Hombre Araña (2002)	139	403
Final Fantasy: The Spirits Within (2001)	137	32
Hulk (2003)	137	132

Nota: las recaudaciones brutas sólo son en Estados Unidos; las recaudaciones mundiales con frecuencia son sustancialmente mayores. Estas cifras no están ajustadas a la inflación.

A propósito. . .

En términos de porcentaje, la película más rentable de todos los tiempos fue *El proyecto de la bruja de Blair*. Producida con \$35 000 por dos amigos que salieron del programa de películas de la Universidad Central de Florida, la película recaudó más de \$140 millones, un ingreso de \$4 000 por cada \$1 que costó producirla.

Solución La figura 7.7 muestra el diagrama de dispersión con costo de producción en el eje horizontal y recaudación bruta en el eje vertical. (Debe verificar algunos de los puntos contra los datos de la tabla 7.2). El diagrama de dispersión se parece mucho al de la figura 7.3g, que

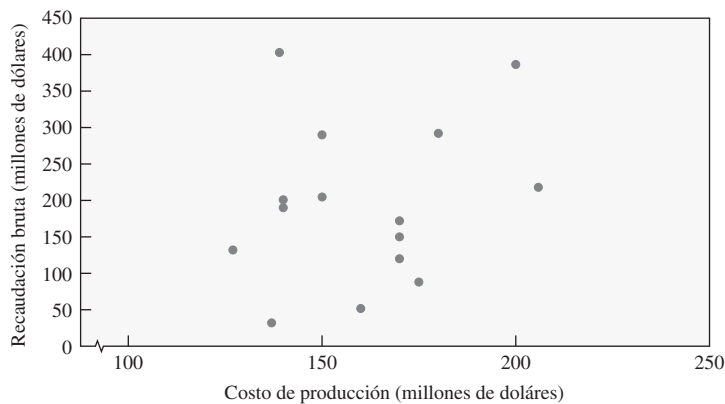


Figura 7.7 Diagrama de dispersión para los datos en la tabla 7.2.

no exhibe correlación, por lo que concluimos que el coeficiente de correlación es cercano a $r = 0$. En otras palabras, al menos para estas películas, no hay correlación entre la cantidad de dinero gastado en producir la película y la cantidad de dinero obtenido en la recaudación bruta.

UN MOMENTO DE REFLEXIÓN

Como práctica adicional, estime visualmente los coeficientes de correlación para los datos de los pesos de los diamantes y sus precios (figura 7.1) y el color del diamante y su precio (figura 7.2).

Cálculo del coeficiente de correlación (sección opcional)

La fórmula para el coeficiente de correlación (lineal) r puede expresarse de varias formas diferentes, todas algebraicamente equivalentes, lo que significa que producen el mismo valor. La expresión siguiente tiene la ventaja de relacionar más directamente la lógica subyacente para r :

$$r = \frac{\sum \left[\frac{(x - \bar{x})}{s_x} \frac{(y - \bar{y})}{s_y} \right]}{n - 1}$$

En la expresión anterior, la división entre $n - 1$ (donde n es el número de pares de datos) muestra que r es un tipo de promedio, por lo que no aumenta simplemente porque se incluyan más pares de datos. El símbolo s_x denota la desviación estándar de los valores x (o los valores de la primera variable) y s_y denota la desviación estándar de los valores y . La expresión $(x - \bar{x})/s_x$ está en el mismo formato que la *puntuación estándar* introducida en la sección 5.2. Usando las puntuaciones estándar para x y y , aseguramos que los valores de r no cambien sólo porque se utilice una escala diferente de valores. La clave para comprender la lógica para r es observar el producto de las puntuaciones estándar para x y las puntuaciones estándar para y . Esos productos tienden a ser positivos cuando existe una correlación positiva, y tienden a ser negativos cuando existe una correlación negativa. Para datos no correlacionados, algunos de los productos son positivos y otros son negativos, con el efecto neto que la suma es relativamente cercana a cero.

La siguiente fórmula alterna para r tiene la ventaja de la simplicidad de los cálculos, de modo que con frecuencia es utilizada siempre que sean necesarios cálculos manuales. La fórmula siguiente también es fácil de incluir en programas estadísticos o en calculadoras:

$$r = \frac{n \times \Sigma(x \times y) - (\Sigma x) \times (\Sigma y)}{\sqrt{n \times (\Sigma x^2) - (\Sigma x)^2} \times \sqrt{n \times (\Sigma y^2) - (\Sigma y)^2}}$$

Esta fórmula es directa de usar, al menos en principio: primero calculamos cada una de las sumas requeridas, luego sustituimos los valores en la fórmula. Asegúrese de observar que (Σx^2) y $(\Sigma x)^2$ no son iguales: (Σx^2) nos pide primero elevar al cuadrado todos los valores de la variable x y luego sumarlos; $(\Sigma x)^2$ nos dice que primero sumemos todos los valores x y luego elevemos al cuadrado esta suma. En otras palabras, realizar primero las operaciones dentro de los paréntesis. De manera análoga, (Σy^2) y $(\Sigma y)^2$ no son iguales.

Sección 7.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Correlación.** ¿Qué es la correlación? ¿El término *correlación* tiene el mismo significado en estadística que en el uso común?
 - Correlación.** Después de calcular el coeficiente de correlación de cinco pares de datos, usted concluye que no existe correlación. ¿Se sigue que una conclusión de 100 pares más de datos similares sería que no habría correlación? ¿Por qué sí o por qué no?
 - Diagrama de dispersión.** Si un coeficiente de correlación se ha calculado como $r = 0.997$, describa el patrón de puntos en el diagrama de dispersión correspondiente.
 - Correlación por coincidencia.** Un estudio mostró una correlación significativa entre el consumo de vino per cápita y el ingreso de maestros. Obvio, es tonto concluir que los maestros gastan sus ingresos adicionales en vino (¿o no es así?) así que identifique al menos otro factor que podría explicar la correlación.
- ¿Tiene sentido?** Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.
- Nacimientos.** Un estudio mostró que para un pueblo, conforme aumentaba la población de cigüeñas, el número de nacimientos también aumentaba. Por tanto, se sigue que el aumento en la población de cigüeñas causa que el número de nacimientos aumente.
 - Efecto positivo.** Un investigador planea investigar la relación entre una disminución en la tasa de mortalidad infantil y el ingreso per cápita de diferentes países. Puesto que una disminución en la tasa de mortalidad infantil es un efecto bueno, o “positivo”, sabemos que existe una correlación positiva entre la tasa de mortalidad infantil y el ingreso per cápita de diferentes países.
 - Correlación.** Dos estudios encontraron una correlación entre bajo peso al nacer y sistemas inmunológicos débiles. El segundo estudio tuvo una muestra mucho más grande, por lo que la correlación encontrada debe ser más fuerte.
 - Interpretación de r .** Un investigador utiliza 20 pares de datos para calcular el valor del coeficiente de correlación lineal y él concluye que $r = 1.2$.

Conceptos y aplicaciones

Tipos de correlación. Los ejercicios 9 al 16 listan pares de variables. Para cada par indique si usted cree que las dos variables están correlacionadas. Si cree que están correlacionadas, indique si la correlación es positiva o negativa. Explique su razonamiento.

- Estatura/peso.** Las estaturas y pesos de 50 mujeres seleccionadas aleatoriamente entre las edades de 1 a 21.
- Estudio/calificaciones.** La cantidad de tiempo que los estudiantes destinan al estudio de un examen de historia y sus calificaciones en ese examen.
- Peso/consumo de combustible.** Los pesos de automóviles y sus tasas de consumo de combustible en millas por galón.
- CI/Tamaño del sombrero.** Las puntuaciones del CI y los tamaños de sombrero de adultos seleccionados aleatoriamente.
- Maratón.** El tiempo (en segundos) que tarda en correr un maratón y el orden de llegada a la meta (primero, segundo, etcétera).
- Altitud/temperatura.** La temperatura exterior en el aire y la altitud de una aeronave.
- Estatura/puntuación del SAT.** Las estaturas y las puntuaciones del SAT de sujetos seleccionados aleatoriamente quienes presentaron el SAT.
- Puntuación del golf/premio monetario.** Puntuaciones del golf y dinero en premio ganado por golfistas profesionales.
- Producción mundial de carne y de grano.** El diagrama de dispersión en la figura 7.8 muestra la relación entre la producción mundial de carne y la producción mundial de grano, ambas medidas en kilogramos por persona. Los puntos de datos corresponden a 11 años diferentes entre 1950 y 2000. Estime el coeficiente de correlación y analice las razones subyacentes para la correlación.

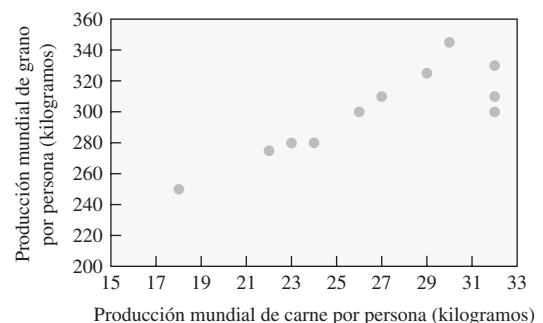


Figura 7.8 Diagrama de dispersión para la información de producción de carne y grano en 11 años diferentes. *Fuente:* Departamento de Agricultura, Estados Unidos.

- Pronóstico a dos días.** La figura 7.9 muestra un diagrama de dispersión en el que la temperatura máxima real para el día es comparada con un pronóstico hecho con dos días de anticipación. Estime el coeficiente de correlación y analice lo que estos datos implican acerca de los pronósticos del clima. ¿Considera que obtendría resultados similares si hiciese diagramas semejantes para otros dos periodos de dos semanas? ¿Por qué sí o por qué no?

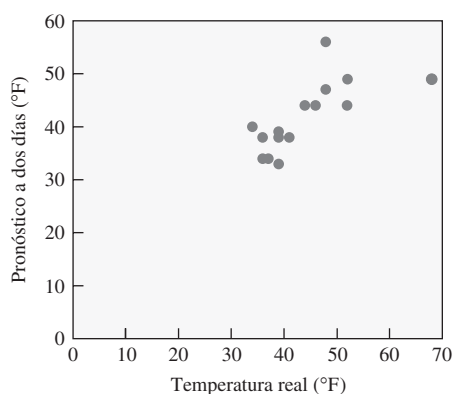


Figura 7.9

19. **¿Velocidades seguras?** Considere la tabla siguiente que muestra los límites de velocidad y las tasas de mortalidad de accidentes automovilísticos en países seleccionados en la década de los ochenta del siglo pasado.

País	Tasa de mortalidad (por cada 100 millones de vehículo-millas)	Límite de velocidad (millas por hora)
Noruega	3.0	55
Estados Unidos	3.3	55
Finlandia	3.4	55
Gran Bretaña	3.5	70
Dinamarca	4.1	55
Canadá	4.3	60
Japón	4.7	55
Australia	4.9	65
Países Bajos	5.1	60
Italia	6.1	75

Fuente: D. J. Rivkin, *New York Times*.

- Construya un diagrama de dispersión.
 - Caracterice brevemente la correlación en palabras (por ejemplo, fuerte correlación positiva, débil correlación negativa) y estime el coeficiente de correlación de los datos. (O calcule el coeficiente de correlación de manera exacta con ayuda de una calculadora o de un programa de cómputo).
 - En el periódico, estos datos fueron presentados en un artículo titulado “El límite de velocidad de cincuenta y cinco mph no es garantía de seguridad”. Con base en los datos, ¿usted coincide con esta afirmación? Explique.
20. **Crecimiento de población.** Considere la tabla siguiente que muestra el cambio porcentual en población y la tasa de natalidad (por cada 1 000 de población) para diez estados durante un periodo de diez años.

Estado	Cambio porcentual en población	Tasa de natalidad
Nevada	50.1%	16.3
California	25.7%	16.9
New Hampshire	20.5%	12.5
Utah	17.9%	21.0
Colorado	14.0%	14.6
Minnesota	7.3%	13.7
Montana	1.6%	12.3
Illinois	0%	15.5
Iowa	−4.7%	13.0
West Virginia	−8.0%	11.4

Fuente: Oficina del Censo de Estados Unidos y Departamento de Salud y Servicios Humanos.

- Construya un diagrama de dispersión para los datos.
 - Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
 - En general, ¿parece que la tasa de natalidad será un buen indicador de la tasa de crecimiento de la población de un estado? Si no, ¿qué otro(s) factor(es) podrían afectar la tasa de crecimiento?
21. **Jugadores más valiosos.** Considere la tabla siguiente que muestra el número de cuadrangulares y el promedio de bateo de los jugadores más valiosos del béisbol (NL = Liga Nacional y AL = Liga Americana)

Jugador	Cuadrangulares	Promedio de bateo
Jeff Kent (2000 NL)	33	.334
Jason Giambi (2000 AL)	43	.333
Barry Bonds (2001 NL)	73	.328
Ichiro Suzuki (2001 AL)	8	.350
Barry Bonds (2002 NL)	46	.370
Miguel Tejada (2002 AL)	34	.308
Barry Bonds (2003 NL)	45	.341
Alex Rodriguez (2003 AL)	47	.298
Barry Bonds (2004 NL)	45	.362
Vladimir Guerrero (2004 AL)	39	.337
Albert Pujols (2005 NL)	41	.330
Alex Rodriguez (2005 AL)	48	.321
Ryan Howard (2006 NL)	58	.313
Justin Moreau (2006 AL)	34	.321

- Construya un diagrama de dispersión para los datos.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- ¿Estos datos sugieren que un alto promedio de bateo es un buen indicador de cuadrangulares? Explique.

22. Datos de películas. Considere la tabla siguiente que muestra la recaudación total en taquilla y la audiencia total para todas las películas estadounidenses, 1990-2002.

Año	Ingresos totales (miles de millones de dólares)	Asistencia total (miles de millones)
1990	5.0	1.18
1991	4.8	1.14
1992	4.9	1.17
1993	5.2	1.24
1994	5.4	1.29
1995	5.5	1.26
1996	5.9	1.34
1997	6.4	1.39
1998	7.0	1.48
1999	7.5	1.47
2000	7.7	1.42
2001	8.4	1.49
2002	9.5	1.64

Fuente: Motion Picture Association of America.

- Construya un diagrama de dispersión para los datos.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- Que el precio de las películas ha aumentado desde 1990, ¿cómo afecta estos datos? Si fuera un directivo de películas, ¿qué conclusiones generales, de esta información, serían importantes para usted?

23. Tiempo de televisión. Considere la tabla siguiente que muestra el número de horas promedio de televisión que se ven en hogares en cinco categorías de ingresos anuales.

Ingreso del hogar	Horas de televisión semanales
Menos de \$30 000	56.3
\$30 000-\$40 000	51.0
\$40 000-\$50 000	50.5
\$50 000-\$60 000	49.7
Más de \$60 000	48.7

Fuente: Nielsen Media Research.

- Construya un diagrama de dispersión para los datos. Para ubicar los puntos utilice el punto medio de cada categoría de ingresos. Utilice un valor de \$25 000 para la categoría “menos de \$30 000” y utilice \$70 000 para “más de \$60 000”.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- Sugiera una razón por qué familias con altos ingresos ven menos televisión. ¿Usted considera que estos datos

implican que puede aumentar sus ingresos simplemente viendo menos televisión? Explique.

24. Clima de enero. Considere la tabla siguiente, que muestra la precipitación mensual media en enero y la temperatura máxima media diaria para diez ciudades del hemisferio norte (Administración Nacional Atmosférica y Oceánica)

Ciudad	Temperatura promedio máxima diaria para enero (°F)	Precipitación media en enero (pulgadas)
Atenas	54	2.2
Bombay	88	0.1
Copenhague	36	1.6
Jerusalén	55	5.1
Londres	44	2.0
Montreal	21	3.8
Oslo	30	1.7
Roma	54	3.3
Tokio	47	1.9
Viena	34	1.5

Fuente: The New York Times Almanac.

- Construya un diagrama de dispersión para los datos.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- ¿Puede sacar conclusiones generales acerca de las temperaturas de enero y las precipitaciones a partir de estos datos? Explique.

25. Ventas al menudeo. Considere la tabla siguiente que muestra las ventas totales de un año (ingreso) y las utilidades para ocho minoristas en Estados Unidos.

Compañía	Ventas totales (miles de millones de dólares)	Utilidades (miles de millones de dólares)
Wal-Mart	315.6	11.2
Kroger	60.6	0.98
Home Depot	81.5	5.8
Costco	60.1	1.1
Target	52.6	2.4
Starbuck's	7.8	0.6
The Gap	16.0	1.1
Best Buy	30.8	1.1

Fuente: Fortune.com

- Construya un diagrama de dispersión para los datos.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- Analice sus observaciones. ¿Volúmenes de venta más altos necesariamente se traducen en mayores ingresos? ¿Por qué sí o por qué no?

26. Calorías y mortalidad infantil. Considere la tabla siguiente que muestra la ingesta de calorías media (todos residentes) y la tasa de mortalidad infantil (por 1 000 nacimientos) para 10 países.

País	Calorías media diarias	Tasa de mortalidad infantil (por cada 1000 nacimientos)
Afganistán	1523	154
Austria	3495	6
Burundi	1941	114
Colombia	2678	24
Etiopía	1610	107
Alemania	3443	6
Liberia	1640	153
Nueva Zelanda	3362	7
Turquía	3429	44
Estados Unidos	3671	7

- Construya un diagrama de dispersión para los datos.
- Caracterice brevemente la correlación en palabras y estime el coeficiente de correlación.
- Analice patrones que observe y cualquier conclusión general a la que pueda llegar.

Propiedades del coeficiente de correlación. Para los ejercicios 27 y 28 determine si la propiedad dada es verdadera y explique su respuesta.

- Intercambio de variables.** El coeficiente de correlación permanece sin cambio si intercambiamos las variables x y y .
- Cambio de unidades de medidas.** El coeficiente de correlación permanece sin cambio si cambiamos las unidades usadas para medir x , y o ambas.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 7 en www.aw.com/bbt.

- Desempleo e inflación.** Utilice la página web de la Oficina de Estadísticas de Trabajo para encontrar tasas de desempleo e inflación durante el año pasado. Construya un diagrama de dispersión para los datos. ¿Observa alguna tendencia?
- Éxito en la NFL.** Encuentre las estadísticas de la temporada de equipos de la NFL. Construya una tabla que muestre lo

siguiente para cada equipo: número de ganados, promedio de yardas ganadas por juego por la ofensiva y yardas promedio por juego permitidas por la defensiva.

Construya diagramas de dispersión para explorar la correlación entre la ofensiva y ganados y la defensiva y ganados. Analice sus hallazgos. ¿Usted considera que existen otras estadísticas del equipo que darían correlaciones más fuertes con el número de ganados?

- Estadística abstracta.** Explore la “tablas de peticiones frecuentes” en el sitio web de *Statistical Abstract of the United States*. Seleccione datos que sean de su interés y explore al menos dos correlaciones. Analice brevemente lo que aprendió de las correlaciones.
- Estatura y largo del brazo.** Seleccione una muestra de por lo menos ocho personas y mida la estatura y largo del brazo de la persona. (Cuando usted mida el largo del brazo, la persona debe estar parada con los brazos extendidos como las alas de un avión). Con los datos muestrales construya un diagrama de dispersión y estime o calcule el valor del coeficiente de correlación. ¿Qué concluye?
- Estatura y pulso.** Seleccione una muestra de por lo menos ocho personas y registre el pulso de cada persona contando el número de latidos en 1 minuto. También registre la estatura de cada persona. Usando las parejas de datos muestrales, construya un diagrama de dispersión y estime o calcule el valor del coeficiente de correlación. ¿Qué concluye?

EN LAS NOTICIAS

- Correlaciones en las noticias.** Determine una noticia reciente que analice algún tipo de correlación. Describa la correlación. ¿El artículo proporciona algún sentido de la fuerza de la correlación? ¿Sugiere que la correlación refleja alguna causalidad subyacente? Analice brevemente si usted cree que las implicaciones hechas por el artículo se deben a alguna correlación.
- Su propia correlación positiva.** Proporcione ejemplos de dos variables que usted espera que estén correlacionadas positivamente. Explique por qué están correlacionadas y por qué la correlación es (o no es) importante.
- Su propia correlación negativa.** Proporcione ejemplos de dos variables que usted espera que estén correlacionadas negativamente. Explique por qué están correlacionadas y por qué la correlación es (o no es) importante.

7.2 Interpretación de correlaciones

Los investigadores examinan a fondo datos estadísticos, de manera constante buscan correlaciones significativas y, con frecuencia, el descubrimiento de una nueva y sorprendente correlación conduce a una gran cantidad de artículos. Podría recordar haber oído acerca de algunas de estas correlaciones descubiertas: el consumo de salvado de avena está correlacionado con el riesgo reducido de cardiopatías; el uso de teléfono celular está correlacionado con el aumento en el riesgo de accidentes automovilísticos; o comer menos está correlacionado con el aumento de la longevidad. Por desgracia, la tarea de *interpretación* de tales correlaciones es mucho más difícil que descubrirlas. Mucho después que los artículos han desaparecido, aún no podríamos asegurar si las correlaciones son significativas y, si lo son, si nos dicen algo de importancia práctica. En esta sección analizamos algunas de las dificultades comunes asociadas con la interpretación de correlaciones.

Las estadísticas muestran que de quienes contraen el hábito de comer, muy pocos sobreviven.

—Wallace Irwin

Cuidado con los valores atípicos

Examine el diagrama de dispersión en la figura 7.10. Su vista quizá le diga que existe una correlación en la que los valores grandes de x tienden a implicar valores grandes de y . En realidad, si calcula el coeficiente de correlación para estos datos, encontrará que es relativamente alto $r = 0.880$, lo que sugiere una correlación muy fuerte.

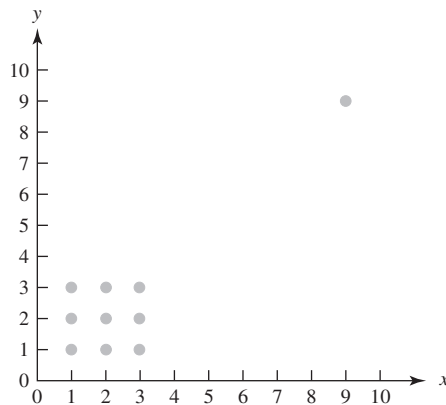


Figura 7.10 ¿Cómo afecta el dato atípico la correlación?

Sin embargo, si coloca su dedo pulgar sobre el punto de datos en la esquina superior derecha de la figura 7.10, la aparente correlación desaparece. De hecho, sin este punto, ¡el coeficiente de correlación es cero! En otras palabras, quitando este punto cambia el coeficiente de correlación de $r = 0.880$ a $r = 0$.

Este ejemplo muestra que las correlaciones pueden ser muy sensibles a los valores atípicos. Recuerde que un *valor atípico* es un valor que es extremo comparado con la mayoría de los otros valores en un conjunto de datos (vea la sección 4.1). Por tanto, debemos examinar los valores atípicos y sus efectos de manera cuidadosa antes de interpretar una correlación. Por un lado, si los valores atípicos son errores en el conjunto de datos, pueden producir correlaciones aparentes que no son reales o esconder la presencia de correlaciones reales. Por otra parte, si los valores atípicos representan datos reales y correctos, nos podrían indicar relaciones que de otra forma serían difíciles de ver.

Observe que aunque debemos examinar los valores atípicos con cuidado, *no* debemos eliminarlos a menos que tengamos una razón poderosa para creer que no pertenecen al conjunto de datos. Incluso en ese caso, los principios de la buena investigación demandan que reportemos los valores atípicos junto con una explicación del porqué consideramos legítimamente eliminarlos.



EJEMPLO 1 Correlación oculta

Usted ha realizado un estudio para determinar cómo el número de calorías que una persona consume en un día está correlacionado con el tiempo que dedica a un enérgico ejercicio de bicicleta. Su muestra consiste en diez mujeres ciclistas, todas de aproximadamente la misma estatura y el mismo peso. Durante un periodo de dos semanas le pidió a cada mujer registrar la cantidad de tiempo que dedican cada día al ciclismo y lo que comieron en cada uno de esos días. Usted utilizó los registros de comida para calcular las calorías consumidas cada día. La figura 7.11 muestra un diagrama de dispersión con el tiempo medio que cada mujer dedicó al ciclismo en el eje horizontal y la ingesta media de calorías en el eje vertical. ¿Mayores tiempos de ciclismo corresponden a mayor ingesta de calorías?

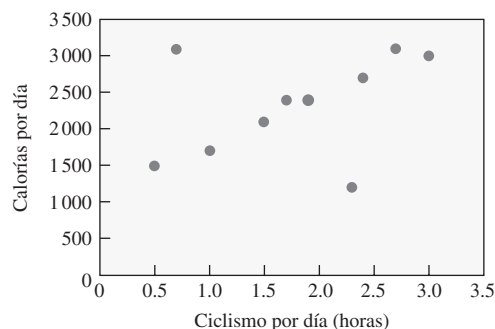


Figura 7.11 Datos del estudio de ciclismo.

Solución Si observa los datos como un todo, su vista le podría decir que existe una correlación positiva en la que a mayor tiempo de ciclismo tiende a ir con mayor ingesta de calorías. Pero la correlación es muy débil, con un coeficiente de correlación de $r = 0.374$. Sin embargo, observe que dos puntos son valores atípicos: uno representa una ciclista que trabaja en la bicicleta alrededor de media hora por día y consume más de 3000 calorías, y el otro representa una ciclista que anda en bicicleta más de dos horas diarias y sólo consume 1200 calorías. Es difícil explicar los dos valores atípicos, dado que todas las mujeres en la muestra tienen estatura y peso similares. Por tanto, podríamos sospechar que estas dos mujeres o registraron datos de manera incorrecta o no siguieron sus hábitos usuales durante las dos semanas de estudio. Si podemos confirmar esta sospecha, entonces tendríamos razón para eliminar los dos datos como inválidos. La figura 7.12 muestra que, sin esos dos valores atípicos, la correlación es muy fuerte y sugiere que el número de calorías consumidas crece en un poco más de 500 calorías por cada hora de ciclismo. Por supuesto, *no* debemos eliminar los valores atípicos sin confirmar nuestra sospecha de que eran puntos de datos no válidos y debemos reportar nuestras razones para dejarlos fuera.

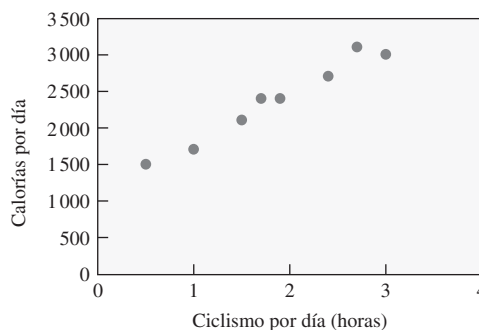


Figura 7.12 Datos de la figura 7.11 sin los dos valores atípicos.

Cuidado con agrupación inadecuada

Las correlaciones también pueden malinterpretarse cuando los datos se agrupan de manera inadecuada. En algunos casos, la agrupación de datos oculta correlaciones. Considere un estudio (hipotético) en el que los investigadores buscan una correlación entre horas de televisión que ven a la semana adolescentes de preparatoria y la calificación promedio escolar (CPE). Los investigadores recolectan 21 parejas de datos en la tabla 7.3.

El diagrama de dispersión (figura 7.13) virtualmente no muestra correlación alguna; el coeficiente de correlación para los datos es de alrededor de $r = -0.063$. La conclusión aparente es que los hábitos de ver la televisión no están relacionados con el aprovechamiento académico. Sin embargo, una investigadora perspicaz se daría cuenta que algunos de los estudiantes ven principalmente programas educativos, mientras que otros tienden a ver comedias, dramas y películas. Por tanto, ella divide el conjunto de datos en dos grupos, uno para los estudiantes que ven principalmente televisión educativa y uno para los otros estudiantes. La tabla 7.4 muestra sus resultados con los estudiantes en estos dos grupos.

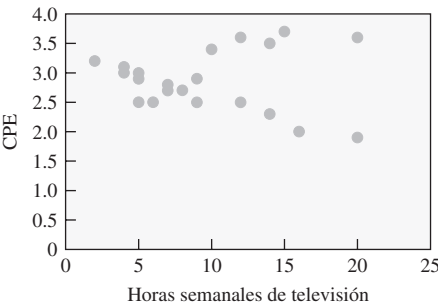


Figura 7.13 El conjunto completo de datos concernientes a horas de televisión y CPE muestra que virtualmente no hay correlación.

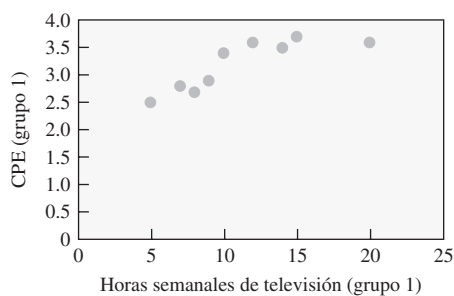
Tabla 7.3 Horas de televisión y CPE en alumnos de preparatoria (datos hipotéticos)	
Hora a la semana de televisión	CPE
2	3.2
4	3.0
4	3.1
5	2.5
5	2.9
5	3.0
6	2.5
7	2.7
7	2.8
8	2.7
9	2.5
9	2.9
10	3.4
12	3.6
12	2.5
14	3.5
14	2.3
15	3.7
16	2.0
20	3.6
20	1.9

Tabla 7.4 Horas de televisión y CPE de preparatoria: datos agrupados (datos hipotéticos)			
Grupo 1: ven programas educativos		Grupo 2: ven programas comunes	
Horas a la semana de televisión	CPE	Horas a la semana de televisión	CPE
5	2.5	2	3.2
7	2.8	4	3.0
8	2.7	4	3.1
9	2.9	5	2.9
10	3.4	5	3.0
12	3.6	6	2.5
14	3.5	7	2.7
15	3.7	9	2.5
20	3.6	12	2.5
		14	2.3
		16	2.0
		20	1.9

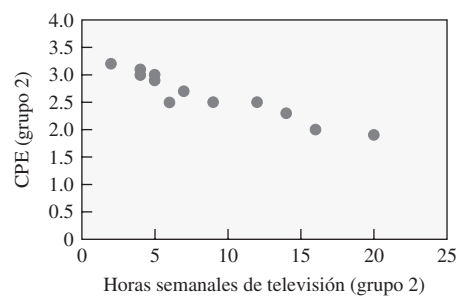
Ahora encontramos dos correlaciones muy fuertes (figura 7.14): una fuerte correlación positiva para los estudiantes que ven programas educativos ($r = 0.855$) y una fuerte correlación negativa para los otros estudiantes ($r = -0.951$). La moraleja de esta historia es que el conjunto de datos originales ocultaron una correlación (hipotética) importante entre televisión y CPE: ver televisión educativa está correlacionado positivamente con la CPE y ver televisión no educativa está correlacionado negativamente con la CPE. Sólo cuando los datos se agruparon de manera apropiada se pudo hacer este descubrimiento.

A propósito...

Niños de 2 a 5 años ven televisión un promedio de 26 horas a la semana, mientras que los niños de 6 a 11 años ven televisión un promedio de 20 horas a la semana (Nielsen Media Research). El tiempo para los adultos es de más de 25 horas a la semana. Si el adulto promedio reemplazase el tiempo de televisión con un trabajo que pague sólo \$8 por hora, su ingreso anual se elevaría en más de \$10 000.



(a)



(b)

Figura 7.14 Estos diagramas de dispersión muestran la misma información que la figura 7.13, separada en los dos grupos identificados en la tabla 7.4.

En otros casos, un conjunto de datos podría mostrar una correlación más fuerte que la que en realidad existe entre subgrupos. Considere los datos (hipotéticos) de la tabla 7.5, recolectados por un grupo consumidor que estudia la relación entre los pesos y los precios de los automóviles. La figura 7.15 muestra el diagrama de dispersión.

Tabla 7.5 Pesos y precios de automóviles

Peso (libras)	Peso (libras)
1500	9500
1600	8000
1700	8200
1750	9500
1800	9200
1800	8700
3000	29000
3500	25000
3700	27000
4000	31000
3600	25000
3200	30000

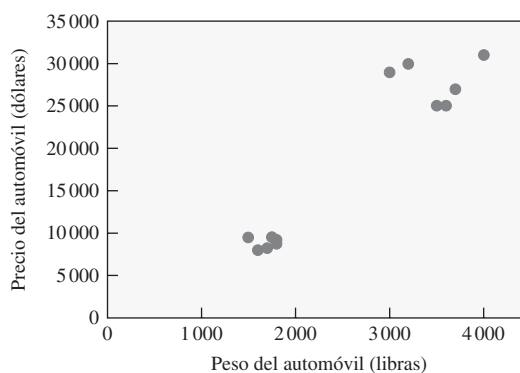


Figura 7.15 Diagrama de dispersión para la información de peso y precio de automóviles de la tabla 7.5.

El conjunto de datos, como un todo, muestra una fuerte correlación; el coeficiente de correlación es $r = 0.949$. Sin embargo, con un examen más cuidadoso, vemos que los datos caen en dos categorías distintas, que corresponden a automóviles ligeros y automóviles pesados. Si analizamos estos subgrupos por separado, ninguno de ellos muestra alguna correlación: los automóviles ligeros, por separado (los primeros seis en la tabla 7.5), tienen un coeficiente de correlación $r = 0.019$ y los automóviles pesados (los seis últimos en la tabla 7.5) tienen un coeficiente de correlación $r = -0.022$. Puede ver el problema examinando la figura 7.15. La aparente correlación del conjunto completo de datos ocurre a consecuencia de la separación entre dos agrupaciones de puntos; no existe correlación dentro de las agrupaciones.

UN MOMENTO DE REFLEXIÓN

Suponga que estuviese comprando un automóvil compacto. Si examina sólo los datos globales y el coeficiente de correlación de la figura 7.15, ¿sería razonable considerar el peso como un factor importante en el precio? ¿Qué sucede si examina los datos por separado para automóviles ligeros y pesados? Explique.

ESTUDIO DE CASO

Búsqueda de correlaciones

El médico Richard Peto de Oxford envió un artículo a la revista médica británica *Lancet*, en el que mostraba que víctimas de ataque al corazón hubiesen tenido mejores posibilidades de sobrevivir si se les hubieran dado aspirinas unas horas después de sus ataques al corazón. Los editores de *Lancet* pidieron a Peto dividir los datos en subconjuntos para ver si los beneficios de la aspirina eran diferentes para diferentes grupos de pacientes. Por ejemplo, ¿la aspirina fue más eficaz para pacientes de cierta edad o para pacientes con ciertos hábitos dietéticos?

Dividir los datos en subconjuntos puede revelar hechos importantes, tales como si los hombres y las mujeres responden de manera diferente al tratamiento. Sin embargo, Peto consideró que los editores le estaban pidiendo dividir la muestra en demasiados subgrupos. Por tanto, él objetó la petición, argumentando que resultaría en correlaciones puramente casuales. Al escribir acerca de esta historia en el *Washington Post*, el periodista Rick Weiss dijo, “Cuando los editores insistieron, Peto se rindió, pero entre otras cosas él dividió a sus pacientes por signo del zodiaco y exigió que sus hallazgos fuesen incluidos en el artículo publicado. Ahora, como un signo de alerta para los inexpertos en estadística, los números jocosos están allí para que todos los vean: la aspirina es inservible para víctimas de ataques cardíacos de los nacidos en Géminis y Libra pero les salva la vida a los nacidos en cualquier otro signo”.

La moraleja de esta historia es que una “expedición de búsqueda” de correlaciones con frecuencia las produce. Lo que no hace que tengan sentido las correlaciones, aunque puedan parecer significativas por medio de medidas estadísticas estándar.

Correlación *no* implica causalidad

Quizá la precaución más importante respecto a la interpretación de las correlaciones es una que ya hemos mencionado: **las correlaciones no necesariamente implican causalidad**. En general, las correlaciones pueden aparecer para cualquiera de las tres razones siguientes:

Posibles explicaciones para una correlación

1. La correlación podría ser una *coincidencia*.
2. Ambas variables correlacionadas podrían ser influidas directamente por alguna *causa subyacente común*.
3. Una de las variables correlacionadas en realidad podría ser una *causa* de la otra. Pero observe que, incluso en este caso, podría ser sólo una de varias causas.

Por ejemplo, la correlación entre la mortalidad infantil y la esperanza de vida en la figura 7.4 es un caso de causa subyacente común: ambas variables responden a la variable subyacente *calidad de cuidado de la salud*. La correlación entre fumar y cáncer de pulmón refleja el hecho que fumar causa cáncer (vea el análisis en la sección 7.4). Correlaciones accidentales también son muy comunes. El ejemplo 2 analiza uno de esos casos.

La precaución acerca de causalidad es particularmente importante en vista de que muchos estudios estadísticos están diseñados para buscar causas. Puesto que estos estudios, por lo general, inician con la búsqueda de correlaciones, es tentador pensar que el trabajo es qué tan pronto se encuentre una correlación. Sin embargo, como analizaremos en la sección 7.4, establecer causalidad puede ser muy difícil.

EJEMPLO 2 Cómo hacerse rico (quizá) en el mercado de valores

Cualquier asesor financiero tiene una estrategia para pronosticar la dirección o tendencia del mercado de valores. La mayoría centrados en datos económicos fundamentales, tales como tasas de interés y utilidades de compañías. Pero una estrategia alterna depende de una correlación sorprendente y bien conocida entre el ganador del Súper Tazón en enero y la dirección del mercado de valores para el resto del año: el mercado de valores tiende a subir cuando un equipo de la vieja liga de 1970, anterior a la NFL, gana el Súper Tazón y tiende a caer cuando el ganador no es de la vieja NFL. Esta correlación coincidió de manera exitosa 31 de los primeros 40 Súper Tazones con el mercado de valores; esta tasa de éxito de 77.5% es más de lo que esperaríamos por el azar. Los ganadores del Súper Tazón XLI en 2007, los Potros de Indianapolis, fueron uno de los equipos de la vieja NFL, de antes de 1970. Con base en este hecho, ¿usted tiene que invertir todos sus ahorros (y quizás algo más que pida prestado) en el mercado de valores después del Súper Tazón de 2007?

Solución Con base en la correlación reportada, podría intentar invertir ya que el ganador de la vieja NFL sugiere un mercado de valores a la alza durante el resto del año. (El equipo que perdió, los Osos de Chicago, también eran de la vieja NFL, por lo que la correlación con

A propósito...

El indicador del Súper Tazón cayó en una mala racha con cuatro años en fila de predicciones incorrectas de 1998 a 2001. Se ha recuperado con predicciones correctas en cuatro de los siguientes cinco años. Este libro se imprimió muy rápido para verificar la predicción de un mercado a la alza en 2007.



el Súper Tazón hubiese pronosticado un mercado a la alza en 2007 sin importar quién ganase). Sin embargo, esta inversión sólo tendría sentido si creyese que el resultado del Súper Tazón en realidad *causa* que el mercado de valores se mueva en una dirección particular. Esto es ridículo y la correlación indudablemente es una coincidencia. Si usted va a invertir no tome como base de su inversión esta correlación.

ESTUDIO DE CASO

Salvado de avena y cardiopatías

Si usted compra un producto que tenga salvado de avena, hay buenas posibilidades de que la etiqueta promocióne los efectos saludables de comer hojuelas de avena. En realidad, varios estudios han encontrado correlaciones en las que la gente que come más salvado de avena tiene menores tasas de cardiopatías. Pero, ¿esto significa que todos debemos comer más hojuelas de avena?

No necesariamente. Sólo porque el consumo de salvado de avena está correlacionado con riesgo reducido de cardiopatías no significa que sea la *causa* de riesgo reducido de cardiopatías. De hecho, en este caso la pregunta de causalidad es muy controversial. Otros estudios sugieren que la gente que come mucho salvado de avena tiende a tener dietas que por lo general son saludables. Así, la correlación entre consumo de salvado de avena y riesgo reducido de cardiopatías puede ser un caso de una causa común subyacente: tener una dieta saludable lleva a la gente tanto a consumir más salvado de avena y a tener un menor riesgo de cardiopatías. En ese caso, para algunas personas, agregar salvado de avena a sus dietas podría ser una *mala* idea ya que podría causarles subir de peso, y el aumento de peso está asociado con el *aumento* en el riesgo de cardiopatías.

Este ejemplo muestra la importancia de tener precaución cuando se consideran temas de correlación y causalidad. Podría pasar mucho tiempo antes de que investigadores médicos sepan con seguridad si el salvado de avena en su dieta en realidad provoca una reducción de riesgo de cardiopatía.

Interpretaciones útiles de correlación

En el estudio de los usos de la correlación que podrían llevar a interpretaciones erróneas, hemos descrito los efectos de valores atípicos, agrupaciones inadecuadas, búsqueda de correlaciones y conclusiones incorrectas de que la correlación implica causalidad. Pero existen muchas interpretaciones correctas y útiles de correlación. Por ejemplo, en la sección 7.1 mostramos cómo la correlación se usa para determinar el precio de diamantes. En otras aplicaciones, la correlación ha sido usada para establecer una relación entre el tamaño de la población y el peso del plástico desechado como basura. Ha sido usada para establecer una relación entre la duración de las erupciones del géiser Viejo Fiel y los intervalos entre las erupciones. En general, la correlación desempeña un papel destacado e importante en una variedad de campos, incluyendo meteorología, investigación médica, negocios, economía, investigación de mercado, psicología y ciencias de la computación.

Sección 7.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Correlación y causalidad.** Al interpretar la correlación es extremadamente importante entender que ésta no implica causalidad. ¿Qué queremos dar a entender cuando decimos que “correlación no implica causalidad”?
- 2. SMRI.** Un artículo en el *New York Times* sobre muertes infantiles afirmó, con base en los resultados del estudio, que si al dormir se coloca a los infantes en posición supina disminuyen las muertes debidas a SMRI (síndrome de muerte repentina infantil). ¿Qué es incorrecto en la afirmación?

- 3. Valores atípicos.** ¿Qué es un valor atípico? Los valores atípicos, ¿cómo podrían afectar las conclusiones acerca de correlación?
- 4. Diagramas de dispersión.** ¿Qué es un diagrama de dispersión y cómo es útil?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Peso y consumo de combustible.** Con base en un estudio que muestra una correlación significativa entre el peso de un automóvil (en libras) y su tasa de consumo de combustible (en millas por galón), podemos concluir que el aumento en el peso del automóvil causa que la tasa de consumo de combustible disminuya.
6. **Ejercicio y salud.** Estudios han mostrado que existe una correlación entre ejercicio y salud. Pero la correlación no implica causalidad. Así, no debemos hacer ejercicio ya que hacerlo no resultaría en mejor salud.
7. **Tomar y conducir.** Puesto que los estudios muestran una correlación entre tomar alcohol y accidentes automovilísticos, podemos concluir que tomar alcohol causa accidentes automovilísticos.
8. **Valor atípico.** Dados 20 pares de datos, suponga que agregamos un par adicional que constituye un valor atípico. El efecto del valor atípico en el coeficiente de correlación será muy pequeño, ya que el valor atípico representa sólo 1 de 21 pares de datos.

Conceptos y aplicaciones

Correlación y causalidad. Los ejercicios 9 al 16 hacen aseveraciones acerca de una correlación. En cada caso indique claramente la correlación. (Por ejemplo, existe una correlación positiva entre la variable A y la variable B). Luego indique si la correlación es más probable debido a una coincidencia, una causa común subyacente o una causa directa. Explique su respuesta.

9. **Armas e índice de crímenes.** En un estado, el número de pistolas registradas ha aumentado constantemente durante los últimos años, y el índice de crímenes también.
10. **Correr y peso.** Se ha encontrado que la gente que hace ejercicio regularmente tiende a pesar menos que las que no corren, y aquellas que corren distancias largas tienden a pesar menos que las que corren distancias más cortas.
11. **Peaje y distancia.** Cuando un conductor viaja más lejos en la autopista de cuota de Massachusetts, el costo del peaje aumenta.
12. **Vehículos y tiempos de espera.** Se ha encontrado que cuando el número de vehículos registrados aumenta, el tiempo que los conductores están parados en el tráfico también aumenta.
13. **Semáforos y accidentes automovilísticos.** Se ha encontrado que cuando el número de semáforos aumenta, el número de accidentes automovilísticos también aumenta.
14. **Galaxias.** Los astrónomos han descubierto que, con la excepción de unas cuantas galaxias cercanas, todas las galaxias en el Universo se mueven alejándose de nosotros. Esto es, entre más distante es una galaxia, mayor es la velocidad a la que se aleja de nosotros.
15. **Gasolina y conducción.** Se ha encontrado que cuando los precios de la gasolina se incrementan, las distancias que los vehículos conducen son más cortas.

16. **Melanoma y latitud.** Algunos estudios han mostrado que, para ciertos grupos étnicos, la incidencia de melanoma (el más peligroso de los cánceres de piel) aumenta cuando la latitud disminuye.

17. **Efectos del valor atípico.** Considere el diagrama de dispersión en la figura 7.16.

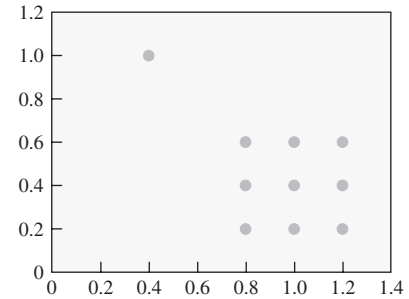


Figura 7.16

- a. ¿Cuál punto es un valor atípico? Ignore el valor atípico, estime o calcule el coeficiente de correlación para los puntos restantes.
- b. Ahora incluya el valor atípico. ¿Cómo afecta el valor atípico al coeficiente de correlación? Estime o calcule el coeficiente de correlación para el conjunto completo de datos.

18. **Efectos del valor atípico.** Considere el diagrama de dispersión en la figura 7.17.

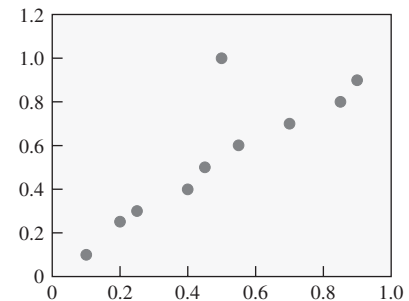


Figura 7.17

- a. ¿Cuál punto es un valor atípico? Ignore el valor atípico, estime o calcule el coeficiente de correlación para los puntos restantes.
- b. Ahora incluya el valor atípico. ¿Cómo afecta al coeficiente de correlación? Estime o calcule el coeficiente de correlación para el conjunto completo de datos.

19. **Datos agrupados de zapatos.** La tabla siguiente proporciona las medidas de pesos y medida del zapato para 10 personas (incluyendo hombres y mujeres).

- a. Construya un diagrama de dispersión para los datos. Estime o calcule el coeficiente de correlación. Con base en este coeficiente de correlación, ¿concluiría que el tamaño del zapato y el peso están correlacionados? Explique.

Peso (libras)	Medida del zapato
105	6
112	4.5
115	6
123	5
135	6
155	10
165	11
170	9
180	10
190	12

- b. Posteriormente, aprendió que los primeros cinco datos en la tabla son de mujeres y los siguientes cinco son de hombres. ¿Cómo cambia esto su visión de la correlación? ¿Aún es razonable concluir que las medidas del zapato y el peso están correlacionadas?
20. **Datos agrupados de temperaturas.** La tabla siguiente muestra la temperatura máxima promedio de enero y la temperatura máxima promedio de julio para 10 ciudades importantes alrededor del mundo.

Ciudad	Máxima de enero	Máxima de julio
Berlín	35	74
Génova	39	77
Kabúl	36	92
Montreal	21	78
Praga	34	74
Auckland	73	56
Buenos Aires	85	57
Sidney	78	60
Santiago	85	59
Melbourne	78	56

- a. Construya un diagrama de dispersión para los datos. Estime o calcule el coeficiente de correlación. Con base en este coeficiente de correlación, concluiría que las temperaturas de enero y julio están correlacionadas para estas ciudades. Explique.
- b. Observe que las primeras cinco ciudades en la tabla son del hemisferio norte y las siguientes cinco son del hemisferio sur. ¿Esto cambia su punto de vista de la correlación? Ahora, ¿concluiría que las temperaturas de enero y julio están correlacionadas para estas ciudades? Explique.
21. **Tasas de natalidad y mortalidad.** La figura 7.18 muestra las tasas de natalidad y mortalidad para diferentes países, medidas en nacimientos y muertes por cada 1 000 habitantes.
- a. Estime el coeficiente de correlación y analice si existe una correlación fuerte entre las variables.
- b. Observe que aparecen en dos grupos de datos dentro del conjunto completo. Haga una conjetura razonable de cómo están conformados estos grupos. ¿En cuál

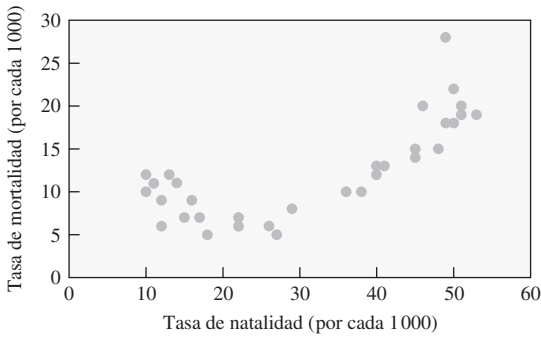


Figura 7.18 Tasas de natalidad y mortalidad para diferentes países. *Fuente:* Naciones Unidas.

- grupo podría encontrar un país relativamente rico como Suecia? ¿En cuál grupo podría encontrar un país pobre como Uganda?
- c. Suponiendo que su conjetura acerca de los grupos en el inciso b es correcta, ¿parece que existe correlación entre los grupos? Explique. ¿Cómo podría confirmar su conjetura acerca de los grupos?
22. **Lectura y calificaciones en exámenes.** El conjunto de datos (hipotéticos) siguiente proporciona el número de horas que leen a la semana 10 estudiantes de sexto grado, y su desempeño en un examen oral estandarizado (máximo de 100).

Tiempo de lectura a la semana	Calificación en el examen oral
1	50
1	65
2	56
3	62
3	65
4	60
5	75
6	50
10	88
12	38

- a. Construya un diagrama de dispersión para estos datos. Estime o calcule el coeficiente de correlación. Con base en este coeficiente de correlación, ¿concluiría que el tiempo de lectura y las calificaciones de exámenes están correlacionados? Explique.
- b. Suponga que sabe que cinco de los niños leen sólo libros de historietas mientras que los otros cinco leen libros regulares. Haga una conjetura de cuáles puntos caen en cada grupo. ¿Cómo podría confirmar su conjetura acerca de los grupos?
- c. Suponiendo que su conjetura acerca de los grupos en el inciso b es correcta, ¿cómo cambiaría su percepción de la correlación entre el tiempo de lectura y calificaciones en exámenes? Explique.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 7 en www.aw.com/bbt.

23. **Actualización de acciones-Súper Tazón.** Encuentre datos de años recientes del ganador del Súper Tazón y el cambio al final del año en el mercado de valores (positivo o negativo). ¿Los resultados recientes aún coinciden con la correlación descrita en el ejemplo 2? Explique.
24. **Correlaciones reales.**
 - a. Describa una situación real en la que exista una correlación positiva que sea resultado de coincidencia.
 - b. Describa una situación real en la que exista una correlación positiva que sea resultado de una causa común subyacente.
 - c. Describa una situación real en la que exista una correlación positiva que sea resultado de una causa directa.
 - d. Describa una situación real en la que exista una correlación negativa que sea resultado de coincidencia.

- e. Describa una situación real en la que exista una correlación negativa que sea resultado de una causa subyacente.
- f. Describa una situación real en la que exista una correlación negativa que sea resultado de una causa directa.

EN LAS NOTICIAS

25. **Correlaciones malinterpretadas.** Encuentre una noticia reciente en la que usted considere que una correlación puede haber sido mal interpretada. Describa la correlación, la interpretación reportada y el problema que ve en la interpretación.
26. **Correlaciones bien interpretadas.** Encuentre una noticia reciente en la que considere que una correlación ha sido presentada con una interpretación razonable. Describa la correlación, la interpretación reportada y explique por qué piensa que la interpretación es válida.

7.3 Rectas de mejor ajuste y pronóstico

Suponga que es muy afortunado y gana un diamante de 1.5 quilates en un concurso. Con base en la correlación entre el peso y el premio de la figura 7.1, debe ser posible pronosticar el valor aproximado del diamante. Sólo necesitamos estudiar cuidadosamente la gráfica y decidir dónde es más probable que se encuentre el punto que corresponde a 1.5 quilates. Para hacer esto, es útil dibujar una **recta de mejor ajuste** (también denominada *recta de regresión*) por los datos, como se muestra en la figura 7.19. La recta es una que “mejor ajusta” en el sentido que, de acuerdo con una medida estadística estándar (que estudiaremos en breve), los puntos están más cercanos a esta recta que a cualquier otra recta que pudiésemos dibujar por los datos.

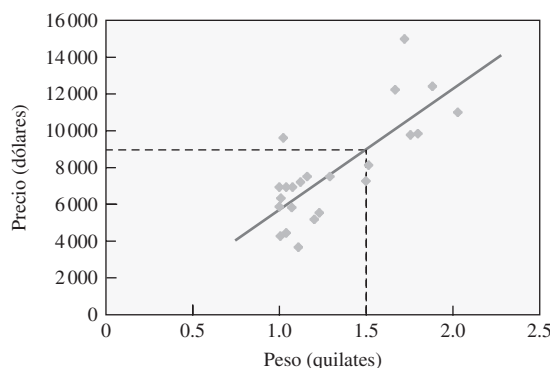


Figura 7.19 Recta de mejor ajuste para los datos de la figura 7.1.

Definición

La **recta de mejor ajuste** (o *recta de regresión*) en un diagrama de dispersión es una recta que está más cercana a los puntos de datos que cualquier otra posible recta (de acuerdo con una medida estadística estándar de cercanía).

A propósito...

El término *regresión* proviene de un estudio de 1877 hecho por sir Francis Galton. Él encontró que las estaturas de los muchachos con padres bajos o altos eran más cercanas a la media que lo que fueron las estaturas de sus padres. Por tanto, él dijo que las estaturas de los hijos *regresaban* (o tenían una *regresión*) hacia la media, de lo cual obtenemos el término *regresión*. Ahora, el término es usado para datos que no tienen nada que ver con una tendencia para regresar hacia una media.



De todas las posibles rectas que pueden dibujarse en un diagrama, ¿cómo sabe cuál es la recta de mejor ajuste? En muchos casos puede hacer una buena estimación de esta recta con sólo observar los datos y dibujar la recta que visualmente parezca que pasa más cercana a todos los puntos de datos. Este método implica dibujar la recta de mejor ajuste “a ojo”. Como puede adivinar, existen métodos para calcular la ecuación precisa de una recta de mejor ajuste (vea el tema opcional al final de esta sección), y muchos programas de cómputo y calculadoras pueden hacer estos cálculos de manera automática. Para nuestros propósitos en este libro, un ajuste a ojo por lo general será suficiente.

Pronósticos con rectas de mejor ajuste

*Es un error capital teorizar
antes de que uno tenga datos.*

—Arthur Conan Doyle

Podemos utilizar la recta de mejor ajuste en la figura 7.19 para predecir el precio de un diamante de 1.5 quilates. Como se indica mediante las líneas discontinuas en la figura, la recta de mejor ajuste predice que el diamante costará alrededor de \$9000. Sin embargo, note que dos puntos reales en la figura corresponden a diamantes de 1.5 quilates, y ambos cuestan menos de \$9000. Así, aunque el precio predicho de \$9000 suena razonable, ciertamente no está garantizado. De hecho, el grado de dispersión entre los puntos en este caso nos dice que *no* debemos confiar en la recta de mejor ajuste para predecir con precisión el precio de cualquier diamante individual. En lugar de eso, la predicción sólo es significativa en un sentido estadístico: nos dice que si examinamos muchos diamantes de 1.5 quilates, su precio medio sería de alrededor de \$9 000.

Ésta es sólo la primera de varias precauciones importantes acerca de la interpretación de predicciones con rectas de mejor ajuste. Una segunda precaución es tener cuidado de utilizar rectas de mejor ajuste para hacer predicciones que vayan más allá de los límites de los datos disponibles. La figura 7.20 muestra una recta de mejor ajuste para la correlación entre mortalidad infantil y longevidad, de la figura 7.4. De acuerdo con esta recta, un país con una esperanza de vida de más de 80 años tendría una tasa de mortalidad infantil *negativa*, lo cual es imposible.

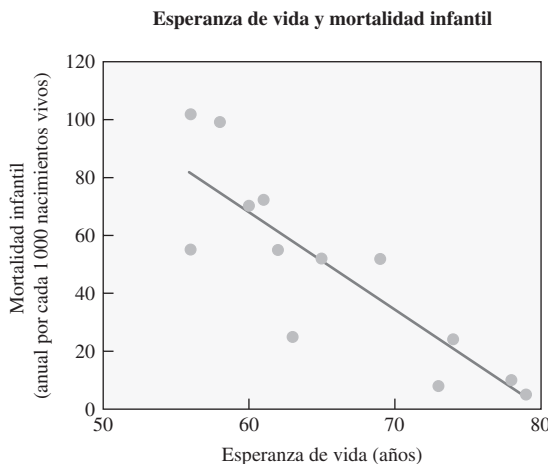


Figura 7.20 Una recta de mejor ajuste para la correlación entre mortalidad infantil y longevidad de la figura 7.4. *Fuente:* Naciones Unidas.

*Es difícil hacer predicciones,
en especial acerca del futuro.*

—Yogi Berra

Una tercera precaución es evitar usar las rectas de mejor ajuste de conjuntos de datos antiguos para hacer predicciones acerca de resultados actuales o futuros. Por ejemplo, economistas que estudian datos históricos encontraron una fuerte correlación negativa entre desempleo y la tasa de inflación. De acuerdo con esta correlación, la inflación debió elevarse dramáticamente en los años recientes, cuando la tasa de desempleo cayó por debajo de 6%. Pero la inflación permaneció baja, lo que muestra que la correlación de datos viejos no continúa cumpliéndose.

Cuarta, una correlación descubierta con una muestra sacada de una población particular, en general, no puede usarse para hacer predicciones acerca de las otras poblaciones. Por ejemplo, no podemos esperar que la correlación entre consumo de aspirina y ataques al corazón en un experimento que incluye sólo a hombres también se aplique a mujeres.

Quinta, recuerde que podemos dibujar una recta de mejor ajuste a través de cualquier conjunto de datos, pero esa recta carece de significado cuando la correlación no es significativa o cuando la relación es no lineal. Por ejemplo, no existe correlación entre el tamaño del pie y el CI, por lo que no utilizaríamos el tamaño del pie para predecir el CI.

Precauciones al hacer predicciones con base en las rectas de mejor ajuste

1. No espere que una recta de mejor ajuste proporcione una buena predicción, a menos que la correlación sea fuerte y existan muchos datos. Si los puntos muestrales están muy cercanos a la recta de mejor ajuste, la correlación es muy fuerte y entonces es más probable que la predicción sea precisa. Si los puntos muestrales están alejados de la recta de mejor ajuste, la correlación es débil y las predicciones tienden a ser mucho menos precisas.
2. No utilice una recta de mejor ajuste para hacer predicciones más allá de los puntos de datos para los cuales la recta se ajustó.
3. Una recta de mejor ajuste con base en datos pasados no necesariamente es válida ahora y podría no resultar en predicciones válidas del futuro.
4. No haga predicciones acerca de una población que sea diferente de la población para la cual se sacó la muestra.
5. Recuerde que una recta de mejor ajuste carece de sentido cuando no hay correlación significativa o cuando la relación no es lineal.

EJEMPLO 1 ¿Predicciones válidas?

Indique si la predicción (o predicción implicada) debe ser confiable en cada uno de los casos, y explique por qué o por qué no.

- a. Encontró una recta de mejor ajuste para una correlación entre el número de horas por día que una persona hace ejercicio y el número de calorías que consume cada día. Ha utilizado esta correlación para predecir que una persona que hace ejercicio 18 horas por día consumirá 15 000 calorías diarias.
- b. Existe una bien conocida, pero débil, correlación entre las calificaciones del SAT y las calificaciones universitarias. Utilice esta correlación para predecir las calificaciones de su mejor amiga a partir de sus calificaciones del SAT.
- c. Datos históricos han mostrado una fuerte correlación negativa entre las tasas de natalidad nacional y la riqueza. Esto es, países con mayor riqueza tienden a tener más bajas tasas de natalidad. Estos datos predicen una tasa alta de natalidad en Rusia.
- d. Un estudio en China ha descubierto correlaciones que son útiles en el diseño de exposiciones en museos que los niños chinos disfrutan. Un director de museo sugiere usar esta información para diseñar una nueva exposición para niños en edad escolar en el área de Atlanta.
- e. Estudios científicos muestran una fuerte correlación entre la ingesta de plomo por niños y retraso mental. Con base en esta correlación, las pinturas que contienen plomo fueron prohibidas.
- f. Con base en un gran conjunto de datos, usted construyó un diagrama de dispersión para el consumo de salsa (por persona) contra los años de educación. El diagrama no muestra correlación significativa, pero de todos modos usted saca una recta de mejor ajuste. La recta predice que alguien que consume una porción de salsa a la semana tiene al menos 13 años de educación.

Apropósito...

En Estados Unidos el plomo fue prohibido en pinturas caseras en 1978 y de latas de alimentos en 1991, y una eliminación progresiva de plomo durante 25 años se terminó en 1995. Sin embargo, un reporte de 1997 de los Centros de Control de Enfermedades aún estimó que un millón de niños menores de 6 años de edad tenían suficiente cantidad de plomo en su sangre para dañar su salud. Las fuentes principales de riesgos de plomo incluyen pintura en casas viejas y la tierra cerca de carreteras importantes, las cuales tienen altos contenidos en plomo por el uso anterior de gasolinas con plomo.

Solución

- Nadie hace ejercicio 18 horas al día de manera continua, por lo que esta cantidad de ejercicio debe estar fuera de los límites de cualquier recolección de datos. Por tanto, una predicción acerca de alguien que haga ejercicio 18 horas al día no debe ser confiable.
- El que la correlación entre las calificaciones del SAT y las calificaciones universitarias sea débil significa que existe mucha dispersión de los datos. Como resultado, no debemos esperar gran precisión si utilizamos esta correlación débil para hacer una predicción acerca de un individuo.
- No podemos suponer de manera automática que los datos históricos aún se aplican ahora. De hecho, Rusia actualmente tiene una muy baja tasa de natalidad, a pesar de tener un bajo nivel de riqueza.
- La sugerencia de utilizar información del estudio chino para una exposición de Atlanta supone que las predicciones hechas de correlaciones en China también se aplican a Atlanta. Sin embargo, dadas las diferencias culturales entre China y Atlanta, la sugerencia del director no debe considerarse sin más información para respaldarla.
- Dada la fuerza de la correlación y la gravedad de las consecuencias, esta predicción y la prohibición que siguió parece muy razonable. De hecho, estudios posteriores establecieron al plomo como una *causa* real de retraso mental, haciendo que las razones atrás de la prohibición sean aún más fuertes.
- Puesto que no hay una correlación significativa, la recta de mejor ajuste y cualesquiera predicciones hechas a partir de ella carecen de significado.

EJEMPLO 2 ¿Las mujeres serán más rápidas que los hombres?

La figura 7.21 muestra los datos y las rectas de mejor ajuste para los tiempos de los récords mundiales para hombres y mujeres en la carrera de una milla. Con base en estos datos, prediga cuándo el récord mundial de las mujeres será más rápido que el récord mundial de los hombres. Comente la predicción.

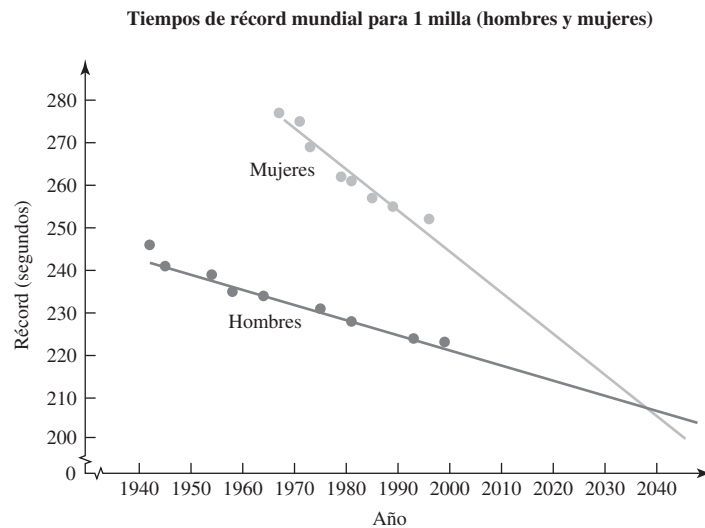


Figura 7.21 Tiempos de récords mundiales en la milla (hombres y mujeres).

Solución Si aceptamos las rectas de mejor ajuste, como las dibujadas, el récord de mujeres será igual al récord de hombres alrededor de 2040. Sin embargo, ésta *no* es una predicción válida ya que está basada en extender las rectas de mejor ajuste más allá del rango de los datos reales. Cierto, no escapa a la pregunta que el récord de mujeres será más rápido que el récord de hombres en 2040. Pero por el mismo razonamiento, ambos récords eventualmente serán cero, implicando que alguien terminará la carrera ¡cuando la pistola de arranque sea disparada!

El coeficiente de correlación y rectas de mejor ajuste

Anteriormente analizamos el coeficiente de correlación como una manera de medir la fuerza de una correlación. También podemos usar el coeficiente de correlación para decir algo acerca de la validez de las predicciones con las rectas de mejor ajuste.

Por razones matemáticas (no estudiadas en este libro), el *cuadrado* del coeficiente de correlación, o r^2 , es la proporción de la variación en una variable que es explicada por la recta de mejor ajuste (o, de manera más técnica, por la relación lineal que la recta de mejor ajuste expresa). Por ejemplo, el coeficiente de correlación para la información del peso y precio de diamantes (vea la figura 7.19) produce $r = 0.777$. Si elevamos al cuadrado este valor, obtenemos $r^2 = 0.604$, que podemos interpretar como sigue: alrededor de 0.6, o 60%, de la variación en los precios de diamante es explicada por la recta de mejor ajuste que relaciona peso y precio. Esto deja 40% de la variación en el precio que puede ser debida a otros factores, presumiblemente a cosas tales como profundidad, mesa, color y claridad —por lo que las predicciones hechas con la recta de mejor ajuste de la figura 7.19 no son muy precisas.

Una recta de mejor ajuste puede dar predicciones precisas sólo en el caso de una correlación perfecta ($r = 1$ o $r = -1$); entonces encontramos $r^2 = 1$, lo que significa que 100% de la variación en la variable puede ser explicado por la recta de mejor ajuste. En este caso especial de $r^2 = 1$, las predicciones deben ser exactamente correctas, excepto por el hecho que la muestra podría no ser una representación verdadera de la información de la población.

Rectas de mejor ajuste y r^2

El *cuadrado* del coeficiente de correlación, o r^2 , es la proporción de la variación en una variable que es explicada por la recta de mejor ajuste.

NOTA TÉCNICA

Con frecuencia los estadísticos llaman a r^2 coeficiente de determinación.

EJEMPLO 3 Contratación para venta

Usted es el administrador de una gran tienda departamental. Al paso de los años ha encontrado una correlación razonablemente fuerte entre sus ventas de septiembre y el número de empleados que necesitará contratar para lograr alta eficiencia durante la temporada de fiestas. El coeficiente de correlación es 0.950. Este año sus ventas de septiembre son bastante fuertes. Con base en la recta de mejor ajuste, ¿debe empezar a colocar anuncios para contratar empleados?

Solución En este caso encontramos que $r^2 = 0.950^2 = 0.903$, lo que significa que 90% de la variación en el número máximo de empleados puede ser explicado mediante una relación lineal con las ventas de septiembre. Esto sólo deja 10% de la variación en el número máximo de empleados sin explicar. Puesto que 90% es muy alto, concluimos que la recta de mejor ajuste explica muy bien la información, por lo que es una buena idea usarla para predecir el número de empleados que necesitará para la temporada de fiestas de este año.

EJEMPLO 4 Concurrencia de votantes y desempleo

Los expertos en ciencias políticas están interesados en conocer qué factores afectan la afluencia de votantes en las elecciones. Uno de tales factores es la tasa de desempleo. Los datos recolectados en años de elección presidencial desde 1964 muestran una muy débil correlación negativa entre la afluencia de votantes y la tasa de desempleo, con un coeficiente de correlación de alrededor de $r = -0.1$ (figura 7.22). Con base en esta correlación, ¿debemos utilizar la tasa de desempleo para predecir la afluencia de votantes en la siguiente elección presidencial?

Solución El cuadrado del coeficiente de correlación es $r^2 = (-0.1)^2 = 0.01$, lo que significa que sólo alrededor de 1% de la variación es explicada por la recta de mejor ajuste. Por tanto, casi toda la variación en los datos debe explicarse por otros factores. Concluimos que el desempleo *no* es un indicador confiable de la afluencia de votantes.

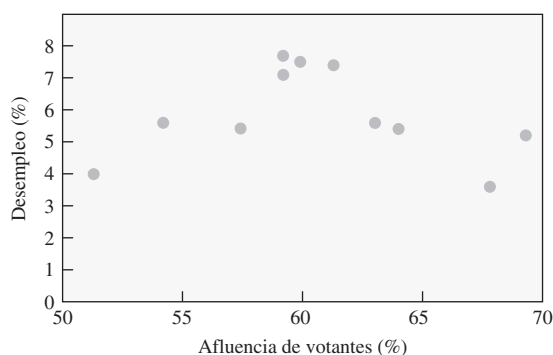


Figura 7.22 Datos de afluencia de votantes y desempleo, 1964-2004. *Fuente:* Oficina de Trabajo, Estados Unidos.

Regresión múltiple

Si en alguna ocasión ha comprado un diamante, podría estar sorprendido de encontrar una correlación débil entre color y precio en la figura 7.2. Seguramente un diamante no puede ser muy valioso si tiene una pobre calidad de color. Quizás el color ayuda a explicar por qué la correlación entre peso y precio no es perfecta. Por ejemplo, quizá las diferencias en color son por lo que dos diamantes con el mismo peso pueden tener precios diferentes. Para comprobar esta idea, sería bueno buscar una correlación entre el precio y alguna combinación de *peso* y *color juntos*.

Todos los que toman su remedio se recuperan en poco tiempo, excepto a los que no los ayuda, quienes mueren. Por tanto, es obvio que falla sólo en casos incurables.

—Galeno, “doctor” romano

UN MOMENTO DE REFLEXIÓN

Compruebe esta idea en la tabla 7.1. Por ejemplo, observe que los diamantes 4 y 5 tienen pesos casi idénticos, pero el diamante 4 cuesta sólo \$4 299, mientras que el diamante 5 cuesta \$9 589. ¿Las diferencias en su color pueden explicar los precios diferentes? Estudie otros ejemplos en la tabla 7.1, en la que dos diamantes tengan pesos similares pero precios diferentes. En general, ¿considera que la correlación con el precio sería más fuerte si utilizamos peso y color juntos, en lugar de uno solo? Explique.

Existe un método para investigar una correlación entre una variable (tal como el precio) y una *combinación* de dos o más variables (tal como peso y color). La técnica se denomina **regresión múltiple**, y en esencia nos permite encontrar una *ecuación de mejor ajuste* que relaciona tres o más variables (en lugar de sólo dos). Puesto que involucra más de dos variables, no podemos hacer diagramas sencillos para mostrar las ecuaciones de mejor ajuste para regresión múltiple. Sin embargo, aún es posible calcular una medida de cuán bien los datos se ajustan a una ecuación lineal. La medida más común de la dispersión en los datos explicada por la ecuación de mejor ajuste es el *coeficiente de determinación*, denotado por R^2 . Nos dice cuánta de la dispersión en los datos es explicada por la ecuación de mejor ajuste. Si R^2 es cercana a 1, la ecuación de mejor ajuste es muy útil para hacer predicciones dentro del rango de los valores de los datos. Si R^2 es cercana a cero, entonces las predicciones con la ecuación de mejor ajuste en esencia no son útiles.

Definición

El uso de **regresión múltiple** nos permite el cálculo de una ecuación de mejor ajuste entre una variable (tal como el precio) y una *combinación* de dos o más variables (tal como peso y color). El coeficiente de determinación, R^2 , nos dice la proporción de dispersión en los datos explicada mediante la ecuación de mejor ajuste.

En este libro no describiremos métodos para determinar las ecuaciones de mejor ajuste mediante regresión múltiple. Sin embargo, podemos utilizar el valor de R^2 para interpretar los resultados de regresión múltiple. Por ejemplo, la correlación entre el precio y *peso y color juntos* da como resultado un valor de $R^2 = 0.79$. Eso es un poco mayor a $r^2 = 0.61$ que

encontramos para la correlación entre el precio y el peso solo. Los estadísticos que estudian la valuación de los diamantes saben que pueden obtener correlaciones más fuertes incluyendo variables adicionales en la regresión múltiple (tal como profundidad, mesa y claridad). Dados los miles de millones de dólares gastados cada año en diamantes, puede estar seguro que los estadísticos desempeñan papeles importantes en la ayuda a los distribuidores de diamantes para obtener las mayores utilidades posibles.

EJEMPLO 5 Contribuciones de alumnos

Ha sido contratado por la Asociación de Alumnos de su colegio para ayudar a estimar cuánto esperaría, de manera razonable, recibir la asociación en una nueva campaña de recolección de fondos. El director sugiere que usted utilice datos concernientes a contribuciones pasadas y nivel de ingresos de alumnos, que tienen un coeficiente de correlación de 0.6. ¿Esta es una buena idea? ¿Puede sugerir una mejor?

Solución El coeficiente de correlación $r = 0.6$ significa que $r^2 = 0.36$, por lo que la relación lineal entre las donaciones y el ingreso de alumnos explica sólo 36% de la variación. Así, usando la recta de mejor ajuste para predecir los montos que donarán de manera individual los alumnos no daría predicciones precisas. Una mejor estrategia sería utilizar una regresión múltiple que incluya factores que podrían influir en las donaciones, tal como áreas de los alumnos, distancia del domicilio actual a la escuela, años desde su graduación y pertenencia a una fraternidad o asociación. Es posible que una ecuación de regresión múltiple bien elegida produzca una correlación mucho más fuerte con las donaciones que cualquier correlación basada en sólo dos variables.

Determinación de ecuaciones para rectas de mejor ajuste (sección opcional)

La técnica matemática para la determinación de la ecuación de una recta de mejor ajuste tiene como base las siguientes ideas básicas. Si dibujamos *cualquier* recta en un diagrama de dispersión, podemos medir la distancia *vertical* entre cada punto y esa recta. Una medida de qué tan bien la recta ajusta los datos es la *suma de los cuadrados* de estas distancias verticales. Una suma grande significa que las distancias verticales de los puntos a la recta son muy grandes y, por tanto, la recta no es un muy buen ajuste. Una suma pequeña significa que los puntos de datos están cerca de la recta y el ajuste es bueno. De todas las posibles rectas, la recta de mejor ajuste es la que minimiza la suma de los cuadrados de las distancias verticales. A consecuencia de esta propiedad, la recta de mejor ajuste en ocasiones se conoce como *recta de mínimos cuadrados*.

Puede recordar que la ecuación de cualquier recta (no vertical) puede escribirse en la forma general

$$y = mx + b$$

donde m es la *pendiente* de la recta y b es la *intersección con el eje y* de la recta. Las fórmulas para la pendiente y la intersección con el eje y de la recta de mejor ajuste son las siguientes:

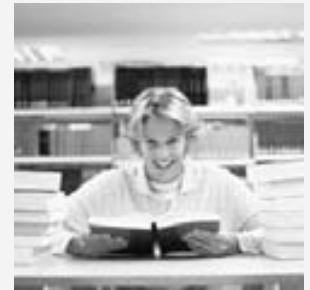
$$\begin{aligned} \text{pendiente} = m &= r \cdot \frac{s_y}{s_x} \\ \text{intersección con el eje y} = b &= \bar{y} - (m \times \bar{x}) \end{aligned}$$

En las expresiones anteriores, r es el coeficiente de correlación, s_x denota la desviación estándar de los valores de x (o los valores de la primera variable), s_y denota la desviación estándar de los valores y , $\bar{x}$ representa la media de los valores de la variable x , y $\bar{y}$ representa la media de los valores de la variable y . Puesto que estas fórmulas son tediosas de usar a mano, por lo regular utilizamos una calculadora o computadora para determinar la pendiente y la intersección con el eje y de las rectas de mejor ajuste.

Cuando los programas de cómputo o una calculadora se utilizan para determinar la pendiente y la intersección de la recta de mejor ajuste, los resultados por lo común se expresan en el formato $y = b_0 + b_1x$, donde b_0 es la intersección con el eje y y b_1 es la pendiente, por lo que sea cuidadoso en identificar de manera correcta esos dos valores.

A propósito...

Un estudio de donaciones de alumnos encontró que, en el desarrollo de una ecuación de regresión múltiple, uno debe incluir estas variables: ingresos, edad, estado civil, si el donante perteneció a una fraternidad o asociación, si el donante está activo en asuntos de alumnos, la distancia del donante al colegio, y la tasa de desempleo nacional, usada como una medida de la economía (Bruggink y Siddiqui, "An Econometric Model of Alumni Giving: A Case Study for a Liberal Arts College". *The American Economist*, volumen 39, núm. 2).



Sección 7.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Recta de mejor ajuste.** ¿Qué es la recta de mejor ajuste (también denominada recta de regresión)? ¿Cómo es útil la recta de mejor ajuste?
- 2. Regresión múltiple.** ¿Qué es una regresión múltiple?
- 3. r^2 .** ¿Qué denota r^2 y cómo puede interpretarse? Esto es, este valor ¿qué nos dice acerca de las variables?
- 4. R^2 .** ¿Qué denota R^2 , y qué nos dice acerca de las variables? ¿Qué nos dice un valor de R^2 cercano a 1?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Valor de r^2 .** Un valor de $r^2 = 0.007$ se obtiene de una muestra de hombres, cada par de datos consisten en la estatura en pulgadas y la estatura en centímetros de un hombre.
- 6. Valor de r^2 .** Un valor de $r^2 = -0.040$ se obtiene de una muestra de hombres, cada par de datos consisten en la estatura en pulgadas y la calificación del SAT para un hombre.
- 7. Estatura y peso.** Con datos de la Encuesta Nacional de Salud, la ecuación de la recta de mejor ajuste para estaturas de mujeres y sus pesos muestra que una mujer de 120 pulgadas de estatura pesa 430 libras.
- 8. El Viejo Fiel.** Con pares de datos muestrales que consisten en el tiempo de duración (en segundos) de erupciones del Viejo Fiel y el intervalo de tiempo (en minutos) después de la erupción, se calculó un valor de $r^2 = 0.926$, indicando que alrededor de 93% de la variación en el intervalo después de la erupción puede explicarse por medio de la relación descrita mediante la recta de mejor ajuste entre esas dos variables.

Conceptos y aplicaciones

Rectas de mejor ajuste sobre diagramas de dispersión. Para los ejercicios del 9 al 12 haga lo siguiente.

- Añada una recta de mejor ajuste al diagrama de dispersión dado.
 - Estime o compare r y r^2 . Con base en su valor para r^2 , determine cuánto de la variación de la variable es explicada por la recta de mejor ajuste.
 - Brevemente analice si podría hacer predicciones válidas con base en esta recta de mejor ajuste.
- Utilice el diagrama de dispersión para el color y el precio en la figura 7.2.
 - Utilice el diagrama de dispersión para la esperanza y la mortalidad infantil en la figura 7.4.

- Utilice el diagrama de dispersión para el número de granjas y el tamaño de las granjas en la figura 7.5.
- Utilice ambos diagramas de dispersión para las temperaturas reales y pronosticadas en la figura 7.6.

Rectas de mejor ajuste. Los ejercicios 13 al 20 se refieren a las tablas de los ejercicios de la sección 7.1. En cada caso, haga lo siguiente.

- Construya un diagrama de dispersión y, con base en una inspección visual, dibuje la recta de mejor ajuste a ojo.
 - Analice brevemente la fuerza de la correlación. Estime o calcule r y r^2 . Con base en su valor para r^2 , diga cuánta de la variación de la variable es explicada por la recta de mejor ajuste.
 - Identifique valores atípicos, si los hay, en el diagrama y analice sus efectos en la fuerza de la correlación y en la recta de mejor ajuste.
 - Para este caso, ¿considera que la recta de mejor ajuste proporciona predicciones confiables fuera del rango de los datos en el diagrama de dispersión? Explique.
- Utilice los datos del ejercicio 19 de la sección 7.1.
 - Utilice los datos del ejercicio 20 de la sección 7.1.
 - Utilice los datos del ejercicio 21 de la sección 7.1.
 - Utilice los datos del ejercicio 22 de la sección 7.1.
 - Utilice los datos del ejercicio 23 de la sección 7.1. Para ubicar los puntos, utilice el punto medio de cada categoría de ingresos, utilice un valor de \$25 000 para la categoría “menos de \$30 000” y utilice un valor de \$70 000 para la categoría “más de \$60 000”.
 - Utilice los datos del ejercicio 24 de la sección 7.1.
 - Utilice los datos del ejercicio 25 de la sección 7.1.
 - Utilice los datos del ejercicio 26 de la sección 7.1.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 7 en www.aw.com/bbt.

- 21. Envenenamiento por plomo.** Investigue envenenamiento por plomo, sus fuentes y sus efectos. Analice las correlaciones que han ayudado a los investigadores a entender el envenenamiento por plomo. Analice los esfuerzos para prevenirlo.
- 22. Asbestos.** Investigue asbestos, sus fuentes y sus efectos. Analice las correlaciones que han ayudado a los investigadores a entender efectos adversos para la salud de la exposición a asbestos. Analice los esfuerzos para prevenir esos efectos adversos para la salud.

23. Indicadores de población mundial. La tabla siguiente proporciona cinco indicadores de la población para once países seleccionados. Estudie estos datos e intente identificar posibles correlaciones. Si es necesario haga investigación adicional, analice las posibles correlaciones que haya encontrado, especule sobre las razones para las correlaciones y analice si sugieren una relación causal. Las tasas de natalidad y de mortalidad son por cada 1 000 habitantes; la tasa de fertilidad es por mujer.

País	Tasa de natalidad	Tasa de mortalidad	Esperanza de vida	Porcentaje urbano	Tasa de fertilidad
Afganistán	50	22	43	20	6.9
Argentina	21	8	72	88	2.6
Australia	15	7	78	85	1.9
Canadá	14	7	78	77	1.6
Egipto	29	8	64	45	3.4
El Salvador	30	6	68	45	3.1
Francia	13	9	78	73	1.6
Israel	21	7	77	91	2.8
Japón	10	7	79	78	1.5
Laos	45	15	51	22	6.7
Estados Unidos	16	9	76	76	2.0

Fuente: *The New York Times Almanac*.

EN LAS NOTICIAS

24. Pronósticos en las noticias. Encuentre una noticia reciente en la que se utilice una correlación para hacer un pronóstico. Evalúe la validez del pronóstico, considere todas las precauciones descritas en esta sección. Globalmente, ¿considera que el pronóstico es válido? ¿Por qué sí o por qué no?

25. Mejor recta de ajuste en las noticias. Aunque en las noticias son raros los diagramas de dispersión, no son desconocidos. Encuentre un diagrama de dispersión de cualquier clase en una noticia (reciente o no). A ojo, dibuje una recta de mejor ajuste. Analice qué pronósticos, si los hay, puede hacer con base en su recta de mejor ajuste.

26. Su propia regresión múltiple. Encuentre un ejemplo de su vida o trabajo en el que un análisis de regresión múltiple podría revelar tendencias importantes. Sin hacer en realidad ningún análisis, en palabras, describa qué buscaría a través de la regresión múltiple y cómo podrían ser útiles las respuestas.

7.4 La búsqueda de la causalidad

Una correlación puede sugerir causalidad, pero por ella misma *nunca* establece causalidad. Para establecer que un factor *causa* otro se requiere mucha más evidencia. Antes, encontramos que una correlación entre dos variables puede ser resultado de (1) coincidencia, (2) una causa subyacente común, o (3) una variable que tenga una influencia directa sobre la otra. El proceso de establecer causalidad es esencialmente un proceso de descartar las primeras dos explicaciones.

En principio, podemos descartar las primeras dos explicaciones mediante experimentos:

- Para descartar coincidencia repetimos el experimento muchas veces o usamos un número grande de sujetos en el experimento. Puesto que la ocurrencia de las coincidencias es aleatoria, no podrán ser consistentes en muchos sujetos o experimentos. Así, aunque una coincidencia podría confundir la situación en unos cuantos ensayos de un experimento, es poco probable que confunda después de muchos ensayos.
- Para eliminar los efectos de las variables de confusión descartemos una causa subyacente común controlando y aleatorizando el experimento (vea la sección 1.3). De esta manera, *sólo* las variables de interés varían entre los grupos de tratamiento y de control. Si los controles descartan las variables de confusión, cualesquiera efectos restantes deben ser causados por las variables que se estudian.

Por desgracia, estas ideas con frecuencia son difíciles de poner en práctica. En el caso de descartar coincidencia, podría consumir demasiado tiempo o ser muy costoso repetir un experimento un número suficiente de veces. Para descartar una causa subyacente común, el experimento debe controlar *todo* excepto las variables de interés, y con frecuencia esto es imposible. Además, existen muchos casos en que los experimentos no son prácticos o no son éticos, por lo que sólo podemos recolectar datos observables. Puesto que estudios observacionales no pueden establecer causalidad de manera definitiva, debemos encontrar otras formas de establecer la causalidad.

Cómo establecer causalidad

Suponga que ha descubierto una correlación y sospecha causalidad. ¿Cómo puede probar su sospecha? Regresemos al tema de fumar y cáncer de pulmón. La fuerte correlación entre fumar y cáncer de pulmón, por ella misma no probó que fumar cause cáncer de pulmón. En principio, podríamos haber buscado una prueba con un experimento controlado. Pero tal experimento sería poco ético, ya que requeriría forzar a un grupo elegido aleatoriamente a fumar cigarrillos. Así que, ¿cómo se estableció que fumar es una causa de cáncer en el pulmón?

La respuesta incluye varias líneas de evidencia. Primero, los investigadores encontraron correlaciones entre fumar y cáncer de pulmón entre muchos grupos de personas: mujeres, hombres y personas de diferentes razas y culturas. Segundo, entre grupos de personas con elementos comunes, el cáncer de pulmón fue muy raro en no fumadores. Tercero, personas que fumaban más y por mayores periodos presentaron altas tasas de cáncer de pulmón. Cuarto, cuando los investigadores explicaban otras causas potenciales de cáncer de pulmón (tales como exposición a gas radón o asbestos), encontraron que casi todos los cánceres de pulmón restantes ocurrieron entre fumadores (o gente expuesta muy de cerca a fumadores).

Estas cuatro líneas de evidencia forman un caso sólido, aunque no descartan la posibilidad de que algún otro factor, tal como genética, predisponga a la gente a fumar y a cáncer de pulmón. Sin embargo, hay dos líneas de evidencia más que hacen improbable el factor genético. Una línea de evidencia provino de experimentos con animales. En experimentos controlados, los animales fueron divididos en grupos de tratamiento y de control elegidos aleatoriamente; los experimentos encontraron una correlación entre inhalar humo de cigarro y cáncer de pulmón, lo cual descarta al factor genético, al menos para los animales. La línea final de evidencia vino de estudios de biología de cultivos de células (hechos con muestras pequeñas de tejido pulmonar humano). Los biólogos encontraron el proceso básico mediante el cual los ingredientes en el humo del cigarro crean mutaciones que causan cáncer. Este proceso no parecía depender en manera alguna de factores genéticos, haciendo casi seguro que el cáncer de pulmón sea causado por fumar y no por algún factor genético preexistente. El hecho de que la exposición al humo de fumadores también esté asociada con algunos casos de cáncer de pulmón descarta más un factor genético (ya que humo de otros fumadores afecta a los fumadores) pero es consistente con la idea de que los ingredientes en el humo del cigarro crean mutaciones causantes del cáncer.

El recuadro siguiente resume estas ideas acerca de cómo se establece la causalidad. En términos generales, el caso para causalidad es más fuerte cuando más de estas directrices se cumplen.

A propósito...

Ante la fuerte evidencia de que fumar causa cáncer, el doctor David Sidransky de la Universidad Johns Hopkins dice: "Tenemos tan sólida prueba molecular que podemos tomar un cáncer individual y, con base en el patrón de cambio genético, determinar si fumar cigarrillos fue la causa de ese cáncer [particular]".

A propósito...

Las primeras cuatro directrices de la derecha son denominadas *Métodos de Mill*, en honor de John Stuart Mill (1806-1873). En su época, Mill fue un líder escolar y uno de los primeros defensores del derecho a votar de las mujeres. En filosofía los cuatro métodos se llaman, respectivamente, los métodos de acuerdo, diferencia, variación concomitante y residuos.

Directrices para establecer causalidad

Si usted sospecha que una variable particular (la causa supuesta) es la causante de algún efecto:

1. Busque situaciones en que el efecto esté correlacionado con la causa supuesta, aunque otros factores varíen.
2. Entre grupos que difieren sólo en la presencia o ausencia de la causa supuesta, compruebe que el efecto está presente o ausente de manera similar.
3. Busque evidencia de que mayores cantidades de causa supuesta producen mayores cantidades del efecto.
4. Si el efecto pudiese ser producido por otras causas potenciales (además de la causa supuesta), asegúrese que el efecto aún permanezca después de descartar estas otras causas potenciales.
5. Si es posible, pruebe la causa supuesta con un experimento. Si el experimento no puede realizarse con humanos por razones éticas, considere hacerlo con animales, cultivos de células o modelos computacionales.
6. Intente determinar el mecanismo físico mediante el cual la causa supuesta produce el efecto.

UN MOMENTO DE REFLEXIÓN

Existe mucha controversia acerca de si los experimentos con animales son éticos. ¿Cuál es su opinión de experimentos con animales? Defienda su opinión.

ESTUDIO DE CASO

Bolsas de aire y niños

A mediados de la década de los noventa, las bolsas de aire del copiloto se hicieron comunes en los automóviles. Estudios estadísticos mostraron que las bolsas de aire salvaban muchas vidas en choques a velocidades de moderadas a altas. Pero también apareció un patrón preocupante. Al menos en algunos casos, niños pequeños, en especial bebés y niños que apenas empiezan a caminar colocados en sillas de automóvil para niños, morían por las bolsas de aire en choques de baja velocidad.

Al principio muchos defensores de las bolsas de aire encontraban difícil creer que éstas pudiesen ser la causa de las muertes. Pero la evidencia observacional se hizo más sólida, cumpliendo las primeras cuatro directrices para establecer causalidad. Por ejemplo, el mayor riesgo para infantes en sillas de automóvil para niños cumplía la directriz 3, ya que indicaba que estando más cerca a las bolsas de aire aumentaba el riesgo de muerte. (Una silla de automóvil para niño se coloca encima del asiento del automóvil, por tanto coloca a un niño más cerca de las bolsas de aire de lo que estaría de otra forma).

Para cerrar el caso, expertos en seguridad realizaron experimentos usando muñecos. Encontraron que los niños, a consecuencia de su tamaño pequeño, estaban sentados donde podían ser fácilmente lastimados por la apertura explosiva de una bolsa de aire. Los experimentos también mostraron que una bolsa de aire podría impactar el asiento de automóvil para niño suficientemente fuerte para causar la muerte, de este modo se reveló el mecanismo físico mediante el cual ocurrían las muertes.

Con el mecanismo físico entendido, era posible desarrollar estrategias para prevenir las muertes. Primero, era obvio que la mejor prevención sería mantener a los niños alejados de las bolsas de aire. Esto llevó a nuevas instrucciones para decir a los padres que las sillas de automóvil para niños *nunca* deben usarse en el asiento delantero, que los niños menores de 12 años deben sentarse en los asientos traseros, si esto es posible. Segundo, los hallazgos condujeron a experimentos adicionales que sugieren que las bolsas de aire podrían proporcionar la misma seguridad con aperturas menos explosivas. Las bolsas de aire que abren de manera menos enérgica ahora son requeridas en todos los automóviles nuevos. Tercero, puesto que el riesgo para niños está relacionado con su tamaño pequeño los adultos pequeños también estarían en riesgo. Esto llevó a las recomendaciones que conductores adultos pequeños (tal como mujeres de baja estatura) se sienten tan lejos de las bolsas de aire como sea posible, y que ellos sólo utilicen automóviles con las más nuevas bolsas de aire que abren de manera menos enérgica.

Causalidad oculta

Hasta ahora hemos estudiado cómo establecer causalidad, después de descubrir una correlación. Sin embargo, en ocasiones las correlaciones —o la carencia de correlación— pueden ocultar una causalidad subyacente. Como lo muestra el siguiente caso de estudio, tal causalidad oculta con frecuencia ocurre a consecuencia de variables de confusión.

ESTUDIO DE CASO

Cirugía de *bypass* para el corazón

La cirugía de *bypass* para el corazón es realizada en gente que tiene obstrucción grave de las arterias que proveen al corazón de sangre (las arterias coronarias). Si el flujo de sangre se detiene en estas arterias, un paciente puede sufrir un ataque al corazón y morir. La cirugía de *bypass* en esencia incluye injertar nuevos vasos sanguíneos en las arterias obstruidas de modo que el flujo de sangre fluya alrededor de las áreas bloqueadas. A mediados de la década de los ochenta muchos doctores estaban convencidos de que la cirugía estaba prolongando las vidas de sus pacientes.

Sin embargo, unos estudios retrospectivos recientes aparecieron con un resultado desconcertante: estadísticamente la cirugía parecía que estaba haciendo poca diferencia. En otras

A propósito...

Muchos estados y países tienen leyes respecto de dónde sentar a los niños en los automóviles. Por ejemplo, en Bélgica es ilegal para un niño menor de 12 años ir sentado en el asiento delantero, si el asiento trasero está disponible.



Apropósito...

Como podrá adivinar, es difícil definir *duda razonable*. Para casos criminales, la Suprema Corte sigue estas directrices de la juez Ruth Bader Ginsburg: "Prueba más allá de una duda razonable es una prueba que lo deja firmemente convencido de la culpabilidad del acusado. Existen muy pocas cosas en este mundo que conozcamos con absoluta certeza, y en casos criminales la ley no requiere pruebas que supere toda duda posible. Si, con base en su consideración de la evidencia, usted está firmemente convencido que el acusado es culpable del crimen del que se le acusa, debe encontrarlo culpable. Si, por otra parte, considera que existe una posibilidad real que él no sea culpable, debe darle el beneficio de la duda y encontrarlo no culpable".



palabras, los pacientes que tenían la cirugía parecían no estar mejor, en promedio, que pacientes similares que no tenían la cirugía. Si esto fuese verdadero, significa que la cirugía no valía la pena por el dolor, riesgo y costo involucrados.

Ya que estos resultados iban en contra de lo que muchos doctores pensaban que ellos habían observado en sus propios pacientes, los investigadores empezaron a ahondar con mayor profundidad. Pronto encontraron variables de confusión que no habían sido tomadas en cuenta en los primeros estudios. Por ejemplo, los pacientes a quienes se les practicaba la cirugía tendían a tener más obstrucciones severas de sus arterias, claro, razón por la cual los doctores recomendaban la cirugía más a estos pacientes. Puesto que estos pacientes estaban en peor forma al inicio, una comparación de la longevidad entre ellos y otros pacientes no era realmente válida.

Más importante, la investigación pronto mostró diferencias sustanciales en los resultados entre pacientes que tenían la intervención quirúrgica en diferentes hospitales. En particular, en pocos hospitales se obtenían éxitos extraordinarios con la cirugía del *bypass* y sus pacientes salían mucho mejor que aquéllos a quienes no se les había aplicado la cirugía o la habían hecho en otros hospitales. Claro, las técnicas quirúrgicas usadas por doctores en los hospitales exitosos fueron por alguna razón diferentes y mejores. Los doctores estudiaron las diferencias para asegurar que todos los doctores podrían estar entrenados en las mejores técnicas.

En resumen, las variables de confusión de *cantidad de obstrucción y técnica quirúrgica* habían evitado que los estudios previos encontrasen una correlación real entre cirugía de *bypass* del corazón y la prolongación de la vida. Ahora, esta cirugía es aceptada como una *causa* de vida prolongada en pacientes con arterias coronarias obstruidas. Ahora está entre los tipos de cirugías más comunes y por lo regular agrega *décadas* de vida a los pacientes que se someten a ella.

Confianza en la causalidad

Las seis directrices nos ofrecen una forma de determinar la solidez de un caso para causalidad, pero con frecuencia debemos tomar decisiones antes que un caso de causalidad esté totalmente establecido. Por ejemplo, considere el bien conocido caso de calentamiento global. Nunca será posible mostrar, más allá de toda duda, que quemar combustibles fósiles es la causa del calentamiento global (vea la sección Hablemos de medio ambiente al final de este capítulo), por lo que debemos decidir si actuar aunque aún nos enfrentemos a alguna incertidumbre acerca de la causalidad. ¿Cuánto más debemos saber antes de decidirnos a actuar?

En otras áreas de estadística hay técnicas aceptadas que nos ayudan a tratar con este tipo de incertidumbre, permitiéndonos calcular un nivel numérico de confianza o de significancia. Pero no existen formas aceptadas de asignar números a la incertidumbre que viene con cuestiones de causalidad. Por fortuna, otra área de estudio ha tratado con problemas prácticos de causalidad por cientos de años: nuestro sistema legal. Podría estar familiarizado con las siguientes tres maneras de expresar un amplio nivel legal de confianza.

Niveles amplios de confianza en causalidad

Causa posible: hemos descubierto una correlación, pero aún no podemos determinar si ésta implica causalidad. En el sistema legal, causa posible (tal como considerar que un sospechoso particular posiblemente cometió un crimen particular) con frecuencia es la razón de iniciar una investigación.

Causa probable: tenemos buena razón para sospechar que la correlación involucra causa, quizá tal vez alguna de las directrices para establecer causalidad se satisfacen. En el sistema legal, causa probable es el estándar general para obtener permiso de un juez para una orden de investigación o intervención telefónica.

Causa más allá de una duda razonable: hemos encontrado un modelo físico que es tan exitoso en explicar cómo una cosa causa otra que parece poco razonable dudar de la causalidad. En el sistema legal, causa más allá de una duda razonable es el estándar usual para condenas y por lo general las demandas que durante el proceso han mostrado cómo y por qué (esencialmente el modelo físico) el sospechoso cometió el crimen. Observe que más allá de *duda razonable* no significa más allá de *toda* duda.

Aunque estos amplios niveles permanecen muy vagos, nos dan al menos un lenguaje común para analizar confianza en causalidad. Si usted estudia leyes, aprenderá mucho más acerca de las sutilezas de la interpretación de estos términos. Sin embargo, ya que la estadística tiene poco que decir acerca de ellas, no las analizaremos mucho más en este libro.

Sección 7.4 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Correlación.** Identifique tres explicaciones diferentes para la presencia de una correlación entre dos variables.
- 2. Papel de los experimentos.** En teoría, podemos utilizar experimentos para excluir dos de las tres diferentes explicaciones para la presencia de una correlación entre dos variables. ¿Cuál de las tres explicaciones *no* queremos excluir? ¿Por qué no queremos descartarla?
- 3. Variables de confusión.** ¿Qué es una variable de confusión? ¿Cómo una variable de confusión puede crear una situación en la que una causalidad subyacente está oculta?
- 4. Correlación y causalidad.** ¿Cuál es la diferencia entre la determinación de una correlación entre dos variables y el establecimiento de causalidad entre dos variables?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- 5. Valor de r .** Si una muestra de datos apareados de dos variables da un coeficiente de correlación “perfecta” de 1, entonces podemos concluir que una de las variables tiene un vínculo causal directo con la otra variable.
- 6. Valor de r .** En un estudio de reacciones adversas provenientes de diferentes cantidades de un tratamiento con fármaco, el coeficiente de correlación se calculó a partir de 20 pares de datos que consisten en las cantidades de fármaco y una medida de la intensidad del dolor en una escala estándar. Si $r = 0.013$, entonces la cantidad de fármaco no puede ser una causa directa única del dolor.
- 7. Fumar y cotinina.** Un estudio mostró que hay una correlación entre la exposición a fumar pasiva y la cantidad medida de cotinina en el cuerpo. Podemos establecer que la exposición a fumar pasiva es una causa de cotinina, si podemos descartar la coincidencia como una posible explicación de la correlación.

- 8. Nicotina.** Suponga que un estudio de parches de nicotina estableció evidencia física de que cuando la nicotina es absorbida por el cuerpo, éste la convierte en cotinina. Se sigue que en un experimento con cinco sujetos, un análisis de la relación entre la cantidad de nicotina proporcionada en los parches y la cantidad de cotinina en el cuerpo mostrará la presencia de una correlación.

Conceptos y aplicaciones

Modelos físicos. Para los ejercicios 9 al 12 determine si la relación causal establecida es válida. Si la relación causal es válida, proporcione una explicación.

- 9. Movimiento de proyectil.** La distancia que una pelota de golf viaja es afectada por la velocidad del palo cuando golpea la pelota.
- 10. Tratamiento con imanes.** Cardiopatías pueden curarse mediante el uso de un brazalete magnético en su muñeca.
- 11. Beber y tiempo de reacción.** Beber mayores cantidades de alcohol disminuye el tiempo de reacción de una persona.
- 12. Altitud y salud.** Cuando la gente escala a altitudes mayores sin oxígeno suplementario, tienden a experimentar aumento de problemas fisiológicos, tales como dolores de cabeza o desorientación.
- 13. Identificación de causas: dolores de cabeza.** Usted está tratando de identificar la causa de sus dolores de cabeza al final de la tarde, que le asedian varios días a la semana. Para cada una de las pruebas y observaciones siguientes, explique cuáles de las seis directrices para establecer causalidad utilizó y qué concluyó. Luego resuma sus conclusiones globales con base en todas las observaciones.
 - a.** Los dolores sólo ocurren en días que usted va a trabajar.
 - b.** Si usted para de tomar Coca en la comida, los dolores de cabeza persisten.
 - c.** En el verano, los dolores de cabeza ocurren con menos frecuencia, si usted abre un poco las ventanas de su oficina.

Ocurren aún menos si abre completamente las ventanas de su oficina.

14. **Fumar y cáncer de pulmón.** Existe una fuerte correlación entre fumar tabaco e incidencia de cáncer de pulmón, y la mayoría de los médicos creen que fumar tabaco provoca cáncer de pulmón. Sin embargo, no todos aquellos que fuman tabaco llegan a desarrollar cáncer de pulmón. Describa brevemente cómo fumar podría causar cáncer cuando no todos los fumadores adquieren cáncer.
15. **Otras causas de cáncer de pulmón.** Además de fumar se ha demostrado que hay otros factores causales probables de cáncer de pulmón. Por ejemplo, la exposición a asbestos y la exposición a gas radón, ambos se encuentran en muchas casas y pueden causar cáncer de pulmón. Suponga que encuentra a una persona que vive en una casa que tiene alto nivel de radón y aislantes que contienen asbesto. La persona le dice, “Yo fumo también, ya que creo que, de todos modos, estoy destinado a tener cáncer de pulmón”. ¿Qué respondería? Explique.
16. **Longevidad de directores de orquesta.** Un famoso estudio en *Forum on Medicine* concluyó que la vida media de los directores de orquestas importantes fue 73.4 años, alrededor de 5 años más que el de todos los estadounidenses en ese tiempo. El autor afirmó que una vida de músico *causa* una vida más larga. Evalúe la afirmación de causalidad y proponga otras explicaciones para la mayor esperanza de vida de los directores.
17. **Madres de mayor edad.** Un estudio presentado en *Nature* asegura que las mujeres que dan a luz a mayor edad tienden a vivir más. De las 78 mujeres que tuvieron al menos 100 años en el momento del estudio, 19% habían dado a luz después de cumplir 40 años. De las 54 mujeres que tenían 73 años en el tiempo del estudio, sólo 5.5% había dado a luz después de los 40 años. Un investigador estableció que “si su sistema reproductor está envejeciendo con la suficiente lentitud, de manera que pueda tener un hijo cuando tenga 40, es un buen presagio de que el resto de usted esté envejeciendo también de manera lenta”. ¿Éste fue un estudio observacional o un experimento? ¿El estudio sugiere que la procreación tardía *causa* mayores tiempos de vida o que tener hijos de manera tardía refleja una causa subyacente? Comente cómo encontró la conclusión del reporte.
18. **Líneas de alto voltaje.** Suponga que las personas que viven cerca de líneas de alto voltaje tienen una mayor incidencia de cáncer que las personas que viven más alejadas de esas líneas. ¿Puede concluir que las líneas de alta tensión son la causa de la elevada tasa de cáncer? Si no, ¿qué otras explicaciones podría haber? ¿Qué otro tipo de investigación le gustaría ver antes de concluir que líneas de alta tensión causan cáncer?

19. **Control de armas.** Con frecuencia, quienes están a favor del control de armas resaltan la correlación positiva entre la disponibilidad de pistolas y tasas de asesinato para apoyar su posición que el control de armas salvaría vidas. ¿Esta correlación, por sí misma, indica que la disponibilidad de pistolas causa una mayor tasa de asesinatos? Sugiera algunos otros factores que podrían apoyar o debilitar esta conclusión.

20. **Vasectomía y cáncer de próstata.** Un artículo titulado “Does Vasectomy Cause Prostate Cancer?” (*Chance*, volumen 10, núm. 1) reporta varios estudios que encontraron un riesgo creciente de cáncer de próstata entre los hombres con vasectomía. En la ausencia de una causa directa, varios investigadores atribuyen la correlación a *sesgo en detección*, en la que hombres con vasectomía son más probables que visiten al doctor y, por tanto, es más probable que el doctor les detecte algún cáncer de próstata. Explique brevemente cómo este sesgo de detección podría afectar la afirmación que la vasectomía causa cáncer de próstata.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 7 en www.aw.com/bbt.

21. **Bolsas de aire y niños.** Empezando en el sitio web de la Administración Nacional de Seguridad en Autopistas, investigue los últimos estudios de bolsas de aire, en especial en consideración con niños. Escriba un breve resumen de sus hallazgos y proporcione recomendaciones para mejorar la seguridad de los niños en los automóviles.
22. **Dieta rica en fibra y enfermedad coronaria del corazón.** En el mayor estudio de cómo una dieta rica en fibra previene la enfermedad coronaria del corazón (CHD, por sus siglas en inglés) en mujeres (*Journal of the American Medical Association*, volumen 281, núm. 21), los investigadores detectaron un riesgo reducido de CHD entre mujeres que tenían una dieta rica en fibra. Encuentre el artículo de investigación, resuma sus hallazgos y analice si una causa para la correlación es la propuesta.
23. **Café y cálculo biliar.** Escrito en el *Journal of the American Medical Association* (volumen 281, núm. 22), investigadores reportaron haber encontrado una correlación negativa entre incidencia de cálculos biliares y consumo de café en hombres. Encuentre el artículo de investigación, resuma los hallazgos y analice si una causa para esta correlación es la propuesta.

- 24. Alcohol y apoplejía.** Investigadores reportaron en *Journal of the American Medical Association* (volumen 281, núm. 1) que el consumo moderado de alcohol está correlacionado con una disminución en el riesgo de apoplejía en personas de 40 años y mayores. (El consumo excesivo de alcohol estuvo correlacionado con efectos nocivos). Encuentre el artículo de investigación, resuma sus hallazgos y analice si una causa para la correlación es la propuesta.
- 25. Tabaco y demandas.** Las compañías de tabaco han sido el tema de muchas demandas relacionadas con los peligros de fumar. Investigue una demanda reciente. ¿Qué trataron de probar los demandantes? ¿Qué evidencia estadística usaron? ¿Cómo considera que establecieron la causalidad? ¿Ganaron? Resuma sus hallazgos en una o dos páginas.

EN LAS NOTICIAS

- 26. Causalidad en las noticias.** Encuentre una noticia reciente en la que una investigación estadística lleve a una conclusión de causalidad. Describa el estudio y la afirmación de causalidad. ¿Considera que la afirmación de causalidad es legítima? Explique.
- 27. Causalidad legal.** Encuentre una noticia concerniente a un caso legal pendiente, ya sea civil o criminal, en el cual establecer causalidad es importante para el resultado. Describa brevemente el tema de causalidad en el caso y cómo la habilidad para establecer o refutar la causalidad influirá en el resultado del caso.

Ejercicios de repaso del capítulo

Para los ejercicios 1 a 4 consulte los datos de cigarros en la tabla. Todas las medidas son en miligramos por cigarro, y todos los cigarros son de 100 milímetros de longitud, con filtro y no son de tipo de mentol o ligeros (con base en datos de la Comisión Federal de Comercio).

Marca	Alquitrán	Nicotina	Monóxido de carbono
Camel	16	1.0	17
Kent	13	1.0	13
Lucky Strike	13	1.1	13
Malibu	15	1.2	15
Marlboro	16	1.2	15
Merit	9	0.7	11
Now	2	0.2	3
Old Gold	18	1.4	18
Pall Mall	15	1.2	15
Winston	16	1.1	18
Camel	16	1.0	17

- El par de datos alquitrán y nicotina tienen un coeficiente de correlación de $r = 0.962$. ¿Qué concluye acerca de la fuerza de la correlación entre alquitrán y nicotina? ¿Qué porcentaje de la variación en nicotina puede ser explicada mediante la correlación entre nicotina y alquitrán?
- El par de datos monóxido de carbono y nicotina tiene un coeficiente de correlación de $r = 0.909$. ¿Qué concluye acerca de la fuerza de la correlación entre monóxido de carbono y nicotina? ¿Qué porcentaje de la variación en nicotina puede ser explicada mediante la correlación entre nicotina y monóxido de carbono?
- El par de datos alquitrán y monóxido de carbono tiene un coeficiente de correlación de $r = 0.979$. ¿Qué concluye acerca de la fuerza de la correlación entre alquitrán y monóxido de carbono? ¿Qué porcentaje de la variación en alquitrán puede ser explicada mediante la correlación entre monóxido de carbono y alquitrán?
- Construya un diagrama de dispersión que utilice los valores dados de alquitrán y nicotina. (Utilice los valores de alquitrán en el eje horizontal). ¿Qué característica de la gráfica indica que hay o no una correlación entre alquitrán y nicotina?
- Con base en un estudio de Suecia, varios periódicos reportaron que “vivir cerca de cables de electricidad causa leucemia en niños”. ¿Cuáles datos son probable que sean la base de tal afirmación? ¿Qué es fundamentalmente incorrecto de la afirmación que la proximidad a cables eléctricos causa leucemia?
- Para diez pares de datos muestrales, el coeficiente de correlación se calculó y es $r = -1$. ¿Qué sabemos acerca del diagrama de dispersión?
- En un estudio de sujetos elegidos aleatoriamente, se encontró que existe una fuerte correlación entre ingresos del hogar y número de visitas al dentista. ¿Es válido concluir que entre mayores sean los ingresos (causa) la gente visite con mayor frecuencia al dentista? ¿Es válido concluir que más visitas al dentista causan que la gente tenga mayores ingresos? ¿Cómo podrían explicarse las correlaciones?
- Usted está considerando la compra más cara que probablemente va a hacer: la compra de una casa. Identifique al menos cinco variables diferentes que es probable afecten el valor actual de una casa. Entre las variables que ha identificado, ¿cuál variable, sola, es probable que tenga la mayor influencia en el valor de la casa? Identifique una variable que es probable que tenga poco o ningún efecto en el valor de una casa.
- Un investigador recolecta pares de datos muestrales y calcula el valor del coeficiente de correlación lineal como 0. Con base en ese valor, él concluye que no existe relación entre las dos variables. ¿Qué es incorrecto en esta conclusión?
- Examine el diagrama de dispersión en la figura 7.23 y estime el valor del coeficiente de correlación.

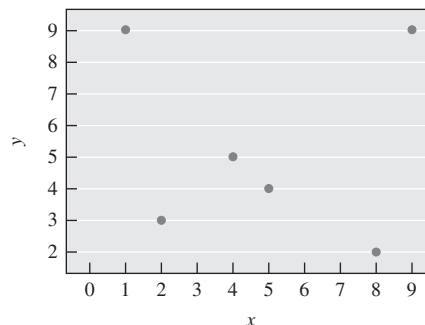


Figura 7.23

Cuestionario del capítulo

1. Llene los espacios en blanco: Todo posible coeficiente de correlación debe estar entre los valores de ____ y ____.
2. ¿Cuál de los siguientes es probable que tengan una correlación?
 - a. Las calificaciones del SAT y los pesos de los sujetos elegidos aleatoriamente.
 - b. Los tiempos de reacción y las puntuaciones de CI de sujetos elegidos aleatoriamente.
 - c. Estatura y extensión de brazos de sujetos elegidos aleatoriamente.
 - d. La proporción de asientos llenados y la cantidad de utilidad de aerolíneas para vuelos elegidos aleatoriamente.
 - e. El valor de automóviles y el ingreso anual de propietarios de automóviles elegidos al azar.
3. Para una colección de parejas de datos muestrales, el coeficiente de correlación fue -0.988 . ¿Cuál de los siguientes enunciados describe la relación entre las dos variables?
 - a. No hay correlación.
 - b. Existe una débil correlación.
 - c. Existe una fuerte correlación.
 - d. Una de las variables es la causa directa de la otra variable.
 - e. Ninguna de las variables es la causa directa de la otra variable.
4. Estime el coeficiente de correlación para los datos en la figura 7.24.
5. Nuevamente consulte el diagrama de dispersión en la figura 7.24. ¿Parece haber una correlación significativa entre las dos variables?

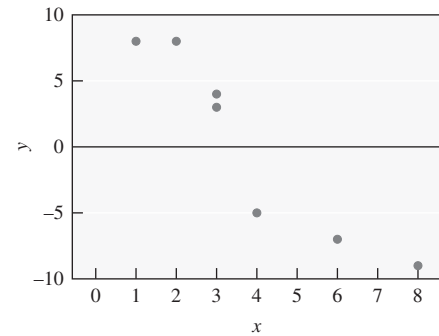


Figura 7.24

En los ejercicios 6 al 10 determine si el enunciado es verdadero o falso.

6. Si existe una fuerte correlación entre dos variables, entonces una de las variables es la causa directa de la otra variable.
7. Un diagrama de dispersión es uno en el que los puntos están dispersos completamente, sin ningún patrón perceptible.
8. Si todos los puntos en un diagrama de dispersión están en una línea recta, iniciando en la parte inferior izquierda y subiendo hacia la parte superior derecha, entonces $r = 1$.
9. Si pares de datos de ingreso e impuesto se recolectan de sujetos elegidos aleatoriamente para los años 1995 a 2008, la presencia de una correlación indica que la relación puede usarse para predecir los impuestos con base en los ingresos en el año 2015.
10. Si $r^2 = 0.09$, entonces es posible que el valor de r sea negativo.

Uso de tecnología

SPSS

Diagrama de dispersión: ingrese pares de datos en columnas en la ventana de edición de SPSS. Dé clic en **Graphs**, luego seleccione el elemento del menú **Scatter/Dot...** Dé clic en la gráfica **Simple Scatter**, luego dé clic en **Define**. En el cuadro de diálogo, dé clic en la variable de la izquierda, y dé clic en el botón de en medio para asignarlo al “eje Y” o al “eje X” como sea apropiado. Dé clic en la segunda variable y asígnela al “eje Y” o al “eje X” como sea apropiado. Dé clic en **OK** y el diagrama de dispersión se mostrará.

Coefficiente de correlación lineal r : ingrese los pares de datos en las columnas de la ventana de edición de datos de SPSS. Dé clic en **Analyze**, seleccione el elemento de menú **Correlate**, luego seleccione **Bivariate**. En el cuadro de diálogo, dé clic en **Pearson**. Dé clic en una variable de la izquierda, y dé clic en el botón central para moverla al cuadro “Variable”. También dé clic en la otra variable, y dé clic en el botón de en medio para moverla al cuadro “Variable”. Dé clic en **OK** para obtener el valor del coeficiente de correlación lineal r , junto con un enunciado acerca de su significancia.

Recta de mejor ajuste: ingrese los pares de datos en columnas en la ventana de edición de datos de SPSS. Dé clic en **Analyze**, seleccione el elemento de menú **Regression**, luego seleccione **Linear**. Dé clic en una variable de la izquierda, y dé clic en el botón de en medio para moverla al cuadro “Dependent” o “Independent(s)” como sea apropiado. Dé clic en **OK** para obtener varias tablas de información. Vea la tabla de “Coefficients” para la intercepción con el eje Y (identificado como “Constant”) y la pendiente de la recta de mejor ajuste.

Excel

Coefficiente de correlación lineal r : primero ingrese los pares de datos muestrales en las columnas A y B. Dé clic en la tecla de función **fx** localizada en la barra principal del menú. Seleccione la categoría de función **Estadística** y el nombre de función **COEF.DE.CORREL**, luego dé clic en **OK**. En el cuadro de diálogo ingrese el rango de celdas para los valores de x , tal como A1:A8. También ingrese el rango de celdas de valores para y , tal como B1:B8. El valor del coeficiente de correlación lineal r se mostrará.

Diagrama de dispersión: Si está utilizando Excel 2003, dé clic en el Asistente de gráficos en el menú principal, luego seleccione el tipo de gráfico identificado como **XY(Dispersión)**. En el cuadro de diálogo ingrese el rango de entrada de los datos, tal como A1:B8. Dé clic en **Siguiente** y proceda a usar los cuadros de diálogo para modificar la gráfica como desee.

Si está usando Excel 2007, utilice el ratón para resaltar los pares de datos. Dé clic en **Insertar** y en el área del gráfico seleccione **Dispersión**. Se mostrarán varias gráficas; dé clic en la de la esquina superior izquierda y se mostrará el diagrama de dispersión. Modifique su gráfica como lo desee.

Recta de mejor ajuste: ingrese los pares de datos en las columnas A y B. Si está usando Excel 2003, seleccione **Herramientas** en el menú principal, luego seleccione **Análisis de datos**, y **Regresión**, luego dé clic en **Aceptar**. Si está usando Excel 2007, dé clic en **Datos**, luego dé clic en **Análisis de datos**, seleccione **Regresión** y luego dé clic en **Aceptar**. Ingrese el rango para los valores de y , tal como B1:B8. Ingrese el rango para los valores x , tal como A1:A8. Dé clic en el cuadro rotulado “Recta de regresión”, luego dé clic en **Aceptar**. Entre toda la información proporcionada por Excel, la pendiente y la intercepción de la ecuación de la recta de mejor ajuste pueden encontrarse bajo la tabla con encabezado “Coeficiente”. La gráfica mostrada incluirá un diagrama de dispersión de los puntos muestrales originales junto con los puntos que serían pronosticados por la recta de mejor ajuste.

STATDISK

Primero ingrese los pares de datos en columnas de la ventana de datos de STATDISK. Seleccione **Analysis** de la barra del menú principal, luego utilice la opción **Correlation and Regression**. Ingrese un valor para el nivel de significancia. Seleccione las columnas que serán usadas. Dé clic en el botón **Evaluate**. La pantalla de STATDISK incluirá el valor del coeficiente de correlación lineal junto con la pendiente e intercepción de la recta de mejor ajuste. Un diagrama de dispersión también se obtendrá dando clic en el botón **Plot**.

HABLEMOS DE EDUCACIÓN

¿Qué ayuda a los niños a aprender a leer?

Todos tienen una idea acerca de la mejor forma de enseñar a leer a los niños. Unos defienden un enfoque fonético, enseñar a los estudiantes cómo “suenan” las palabras. Algunos defienden un enfoque de “lenguaje integral”, enseñar a los estudiantes a reconocer palabras de su contexto. Otros abogan por una combinación de estos enfoques o algo completamente diferente. Estas discrepancias de ideas no sería importante si sólo fuesen opiniones. Pero en un país que gasta aproximadamente un *billón de dólares* al año en educación, diferentes enfoques para enseñar a leer implica importantes confrontaciones políticas entre grupos con diferentes intereses especiales. Un cambio en la política puede causar un repentino cambio en las políticas escolares. Por ejemplo, en 1998 la asamblea legislativa de California aprobó leyes que hacían los fondos escolares dependientes del movimiento de la escuela hacia un enfoque de lenguaje integral para la lectura.

Las enormes apuestas involucradas en la enseñanza de la lectura demandaron estadísticas para medir la efectividad de distintos enfoques. Al menos políticamente las estadísticas educativas más importantes son aquellas que provienen de la NAEP (por sus siglas en inglés de Valoración Nacional de Progreso Educacional), con frecuencia simplemente conocido como “Reporte de la Nación”. La NAEP es una encuesta continua de valoración estudiantil llevada a cabo por una agencia gubernamental, Centro Nacional para Estadísticas Educativas, con autorización y fondos del Congreso de Estados Unidos.

La NAEP utiliza muestreo aleatorio estratificado (vea el capítulo 1) para elegir muestras representativas de estudiantes de cuarto, octavo y décimo segundo grados con variedad de grupos étnicos, ingresos familiares, tipo de atención escolar, etcétera. A los estudiantes elegidos para las muestras se les aplican exámenes diseñados para medir aprovechamiento académico en un área particular, tal como lectura, matemáticas o historia. Las muestras se eligieron a nivel nacional y estatal. Globalmente, se eligieron unos cuantos miles de estudiantes para cada examen. Los resultados de los exámenes de la NAEP de manera inevitable son noticia, con artículos que ponderan o condenan caídas en las calificaciones del examen. También tienen impacto político. Por ejemplo, el movimiento de California hacia el lenguaje integral ocurrió después de que su clasificación en los exámenes de la NAEP fue el lugar 45 entre los 50 estados.

Pero, ¿qué es lo que realmente causa una mejora en el desempeño de la lectura? Los investigadores inician por investigar correlaciones entre desempeño de lectura y otros factores. Algunas correlaciones son claras, pero no ofrecen camino para mejorar la lectura. Por ejemplo, la educación de los padres está claramente correlacionada con los logros de lectura: los niños con más alto nivel educativo de los padres tienden a leer de manera más competente que aquellos con padres sin educación. Pero esta correlación no ofrece mucha guía para las escuelas, ya que los niños no pueden reemplazar a sus padres. Otras veces las correlaciones pueden sugerir formas de mejorar la lectura. Por ejemplo, los estudiantes que reportan leer más páginas por día en la escuela y para las tareas, tienden a tener calificaciones más altas que los que leen menos páginas. Esto sugiere que las escuelas deben asignar más lectura.

Por supuesto, la gran apuesta involucrada en la educación hace de las estadísticas educativas particularmente propensas a malas interpretaciones o mal uso. Considere unos de los problemas que hacen los exámenes de lectura de la NAEP difíciles de interpretar:

- Son “exámenes estandarizados” que son los mismos para todos los estudiantes examinados y tienden a ser, en su mayor parte, de opción múltiple. Algunas personas creen que tales exámenes de manera inevitable están sesgados y no pueden medir en realidad la habilidad de lectura.



- Puesto que los exámenes, por lo general, no afectan las calificaciones de los estudiantes, algunos estudiantes no los toman en serio, en cuyo caso los exámenes no reflejan la capacidad real de lectura. Por ejemplo, algunos administradores de examen han reportado estudiantes que hacen dibujos con los “ovalitos” de los exámenes de opción múltiple, en lugar de tratar de responder las preguntas con su mejor capacidad.
- Las comparaciones entre estados podrían no ser válidas, si la conformación de la población estudiantil varía de manera significativa entre los estados. Por ejemplo, California tiene un alto porcentaje de estudiantes para quienes el inglés es un segundo idioma. Algunas personas creen que esto explicaba las bajas calificaciones del estado, en lugar de algo que hacer con las técnicas de enseñanza.
- Hay alguna evidencia de trampas en la parte de los *adultos* involucrados en los exámenes de la NAEP, por ejemplo, la elección de muestras que en realidad no son representativas sino sesgadas hacia los estudiantes que leen mejor. Esta trampa puede ser motivada por el hecho que las escuelas, los distritos escolares y estados se clasifican de acuerdo con los resultados de la NAEP. Altas calificaciones pueden llevar a recompensas para maestros y administradores en la forma de aumento en los fondos o salarios mayores, mientras que calificaciones bajas pueden provocar diversas acciones de castigo.

Usted probablemente piense en otros problemas que hacen difícil interpretar los resultados de la NAEP. Así, no debe sorprenderse de que la lectura continúe siendo un enorme campo político de batalla. Así, ¿qué puede hacer, como individuo, para ayudar a un niño a leer? Por fortuna, los estudios de la NAEP también revelan algunas correlaciones que no son controversiales y coinciden con el sentido común. Por ejemplo, desempeños altos de lectura están correlacionados con cada uno de los factores siguientes:

- más lectura total, tanto para la escuela y por placer
- más opciones de lectura —esto es, permitir que los niños seleccionen sus propios libros para leer
- más escritura, en particular ampliar las partes tales como ensayos o cartas extensas
- más discusión del material de lectura con amigos y la familia
- menos horas viendo la televisión

Estas correlaciones dan al menos una guía de cómo ayudar a aprender a los niños a leer y deben ser buenos puntos de inicio para análisis de cómo aumentar la alfabetización. Por supuesto, los políticos y grupos con interés especial probablemente encontrarán formas para hacer que estos resultados se ajusten, sin importar lo que ellos pudiesen tener en sus agendas preconcebidas. Así, si usted tiene opiniones sólidas acerca de técnicas de enseñanza, puede unirse a las batallas que probablemente continuarán durante las siguientes décadas.

PREGUNTAS PARA DISCUSIÓN

1. Un resultado claro de los exámenes de lectura de la NAEP es que los estudiantes de escuelas privadas tienden a tener calificaciones significativamente mayores que los estudiantes de escuelas públicas. ¿Esto implica que las escuelas privadas son “mejores” que las escuelas públicas? Defienda su opinión.
2. ¿Usted considera que los exámenes estandarizados como los de la NAEP son formas válidas de medir desempeño académico? ¿Por qué sí o por qué no?
3. Actualmente, los exámenes de la NAEP son aplicados a unos cuantos miles de los millones de niños en edad escolar en Estados Unidos. Algunas personas defienden la aplicación de exámenes similares a todos los estudiantes, ya sea de manera voluntaria u obligatoria. ¿Considera que tales “exámenes nacionales estandarizados” son una buena idea? ¿Por qué sí o por qué no?
4. Una correlación que aún no ha sido estudiada cuidadosamente es la correlación entre el uso de la computadora y la lectura. ¿Considera que el uso de una computadora e internet ayuda o perjudica a los niños en términos de aprendizaje de lectura? ¿Por qué?

5. Lea la última edición de la NAEP (disponible en línea, el vínculo aparece abajo). ¿Cuáles son algunos de los últimos resultados en consideración a la enseñanza de la lectura en Estados Unidos?

LECTURAS RECOMENDADAS

Lemann, Nicholas, “The Reading Wars”, *The Atlantic Monthly*, noviembre de 1997.

NAEP Reading Report Card, la última edición está en línea en <http://nces.ed.gov/nationsreportcard/reading>.

Schemo, Diana, “In War over Teaching Reading, a U.S.-Local Clash”, *New York Times*, 9 de marzo de 2007.

HABLEMOS DE MEDIO AMBIENTE



¿Qué causa el calentamiento global?

La sección Hablemos de medio ambiente del capítulo 3 (página 140) presentó datos visuales del caso en que el calentamiento global está ocurriendo. Aquí ponemos nuestra atención a una pregunta más profunda: ¿el calentamiento global observado está ocurriendo por razones naturales, o lo estamos causando por nuestro uso de combustibles fósiles y otras actividades? Esta pregunta tiene muchas ramificaciones políticas y económicas. Si el calentamiento es natural, entonces no puede ser causa de alarma, ya que de manera natural podría revertirse o podríamos hacer poco. Pero si el calentamiento está siendo causado por nosotros, entonces si se continúa actuando como lo hemos hecho en el pasado se exacerbará el calentamiento, en cuyo caso es imperativo emprender cambios políticos y económicos para aliviar el problema antes de amenazar nuestra civilización.

Actualmente existe un consenso claro entre los científicos que estudian el clima de la Tierra en favor del punto de vista que los humanos somos la causa del calentamiento global. En 2007 el Panel Intergubernamental sobre el Cambio Climático (IPCC) —un panel de más de 600 científicos del clima de todo el mundo— concluyó que hay al menos 90% de certidumbre que el calentamiento observado durante los últimos 50 años ha sido causado por las emisiones humanas de gases tales como dióxido de carbono, y que emisiones continuas calentarían el planeta al menos varios grados más durante este siglo, causando cambios sustanciales a los patrones de clima y elevando el nivel del mar. ¿Cómo llegaron los científicos a esta conclusión?

Para entender los argumentos, primero debemos revisar la ciencia básica subyacente del calentamiento global: el **efecto invernadero**, que atrapa el calor, causado por ciertos gases atmosféricos. El mecanismo físico del efecto invernadero es sencillo y ha sido comprendido durante muchas décadas. Actúa así (figura 7.25): la luz solar es una forma de energía que calienta la superficie de un planeta. Para que la temperatura de la superficie permanezca estable, esta energía absorbida debe ser regresada al espacio (de otra forma, la absorción continua de luz solar rápidamente elevaría la temperatura en una cantidad enorme), y las superficies planetarias regresan esta energía en la forma de luz infrarroja. (La luz infrarroja no la puede ver con sus ojos, pero es fácilmente detectada con cámaras). El efecto invernadero ocurre cuando gases como vapor de agua (H_2O), dióxido de carbono (CO_2) y metano (CH_4) —llamados **gases de invernadero**— absorben parte de esta luz infrarroja en su camino hacia arriba. Estos gases reemiten la luz infrarroja en direcciones aleatorias, por lo que su efecto global es mantener más energía concentrada cerca de la superficie del planeta, haciéndola más caliente de lo que sería de otra forma. Entre mayor sea la abundancia de gases de invernadero, mayor será la detención de energía y el planeta se vuelva más caliente.

Observe que aunque aún existe un poco de incertidumbre acerca del nivel de calentamiento global inducido por los humanos, no hay duda científica acerca del mecanismo del efecto invernadero. En realidad, el efecto invernadero ocurre de manera natural en la Tierra, y que se presente es algo muy bueno. Sin el efecto invernadero, la temperatura global de la Tierra sería cercana a 0°F , haciendo demasiado frío nuestro planeta para agua líquida o para la vida. A consecuencia de la ocurrencia natural del efecto invernadero, la temperatura promedio global real es cercana a 60°F . Estudios de otros planetas confirman nuestra comprensión del efecto invernadero, y muestra que puede ser todavía más importante que la distancia del planeta al sol, en la determinación de la temperatura de la superficie del planeta. Venus proporciona el ejemplo más extremo: aunque Venus es más cercano al Sol que la Tierra, sus nubes son tan reflectoras que su superficie recibe menos luz solar que la superficie de la Tierra. A consecuencia

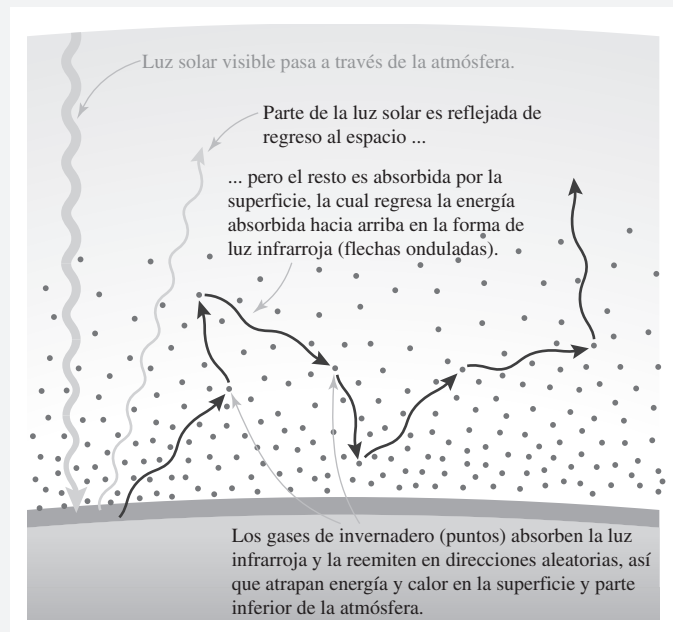


Figura 7.25 Este diagrama muestra el mecanismo básico del efecto invernadero. Entre mayor sea la abundancia de gases de invernadero, se hace más lento el escape de luz infrarroja y se provoca el calentamiento del planeta.

de esto, sin el efecto invernadero, Venus sería más frío que la Tierra. Pero puesto que Venus tiene una atmósfera gruesa que contiene mucho más dióxido de carbono que la atmósfera de la Tierra (por un factor de alrededor de 170 000), la temperatura real de la superficie es abrasadora, 870°F. Dado que el efecto invernadero ocurre de manera natural podemos concluir que es algo bueno para la vida en la Tierra; por otro lado, Venus muestra lo que es tener demasiado de una cosa buena.

Es esta certeza científica acerca del mecanismo del efecto invernadero que condujo a la preocupación por la posibilidad del calentamiento global provocado por el hombre. Ya que la actividad humana, como la quema de combustibles fósiles (junto con otras actividades como la tala o quema de bosques), libera dióxido de carbono y otros gases de invernadero en la atmósfera, podríamos esperar un aumento en la concentración de gas de invernadero atmosférico, lo cual a su vez calentaría nuestro planeta más allá de su nivel natural de calentamiento de invernadero. En realidad, no hay duda de que si seguimos agregando de manera indefinida gases a la atmósfera, al final causaremos suficiente calentamiento global para poner en peligro nuestra supervivencia. Las preguntas son si (y cuánto) el calentamiento ya está en camino y si podemos separar cambios en la concentración de gas invernadero debido a actividad humana de cambios que podrían deberse a factores naturales.

En un intento por responder estas preguntas, Estados Unidos y otros países han dedicado miles de millones de dólares durante las últimas dos décadas a un esfuerzo sin precedentes para entender el clima de la Tierra. Esta investigación ha sido usada para presentar el caso de que la generación de gases de invernadero por parte de los humanos es la causa del calentamiento global. Tres líneas de evidencia hacen el caso particularmente fuerte.

Primera, como se estudió en el capítulo 3, ahora los datos muestran una tendencia clara al calentamiento en el último siglo, con el calentamiento acelerándose en las últimas tres décadas (vea la figura 3.60, página 141). Estos datos han terminado de manera efectiva con cualquier debate acerca de si el calentamiento está en camino, dejándonos por saber si el calentamiento es debido a factores humanos o a factores naturales.

La segunda línea de evidencia proviene de mediciones cuidadosas de concentraciones de dióxido de carbono en el pasado y en el presente en la atmósfera terrestre. Si revisa estos datos,

mostrados en la figura 3.61 (página 141), puede ver que cambios pasados en la concentración de dióxido de carbono están correlacionados claramente con cambios en la temperatura, confirmando nuestra esperanza que una concentración mayor de gases de invernadero va acompañada con temperaturas más altas. (Sin embargo, estos datos no establecen por sí mismos que los gases de invernadero *causen* temperaturas más altas; esa comprensión viene del mecanismo físico). Observe que aunque datos pasados muestran que la concentración de dióxido de carbono varía de manera natural, también muestran que el aumento reciente es mucho mayor que cualquier incremento natural de los últimos cientos de miles de años. La actividad humana es la única explicación viable para el enorme aumento reciente en la concentración de dióxido de carbono.

La tercera línea de evidencia proviene de experimentos. No podemos realizar experimentos controlados con todo nuestro planeta, pero podemos correr experimentos con *modelos de computadora* que simulan la manera en que funciona el clima de la Tierra. Éste es increíblemente complejo, y quedan muchas interrogantes en los intentos por modelar el clima en computadoras. Sin embargo, los modelos actuales son el resultado de décadas de trabajo y refinamientos. Cada vez que un modelo del pasado falla en concordar con datos reales, los científicos buscan entender los ingredientes faltantes (o incorrectos) en el modelo y luego tratan de nuevo con modelos mejorados. Los modelos actuales no son perfectos, pero coinciden bien con datos de clima actuales, lo que genera en los científicos expectativas de que los modelos tienen valores proféticos. La figura 7.26 compara datos de modelos con datos reales. Lo más importante, los modelos coinciden sólo cuando incluyen los efectos de los gases invernadero puestos por los humanos en la atmósfera. Los modelos que sólo consideran factores naturales fallan en coincidir con los datos observados.

La conclusión es clara. Con base en los modelos, los datos disponibles y nuestra comprensión del mecanismo del efecto invernadero, podemos tener más certeza de que el calentamiento global es causado principalmente por la liberación humana de dióxido de carbono y otros gases de invernadero.

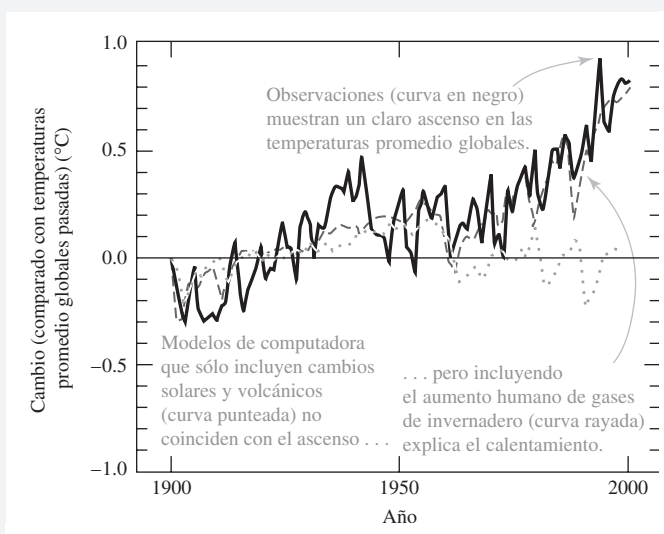


Figura 7.26 Esta gráfica compara los cambios de temperatura observada (curva en negro) con las predicciones de los modelos para el clima. La curva punteada representa las predicciones del modelo que sólo incluye factores naturales, tales como cambios en la brillantez del Sol y efectos de volcanes. La curva rayada representa las predicciones del modelo que incluye la contribución humana debida al incremento de concentración de gases de invernadero junto con factores naturales. Observe que sólo la curva rayada coincide bien con las observaciones, en especial en décadas recientes, proporcionando una fuerte evidencia de que el calentamiento global es un resultado de la actividad humana. (Cada una de las curvas rayada y punteada para los modelos son promedios de muchos modelos independientes de científicos para el calentamiento global, lo cuales por lo general coinciden uno con otro en $0.1^{\circ}\text{C} - 0.2^{\circ}\text{C}$).

PREGUNTAS PARA DISCUSIÓN

1. Revise las seis directrices para establecer causalidad en la página 316. Analice si y cómo cada directriz se cumple para datos actuales y para comprender el calentamiento global.
2. El IPCC concluyó que el calentamiento global es “muy probablemente” causado por nosotros, pero aún hay algunas personas que argumentan lo contrario. Haga una búsqueda en la web para encontrar algunos de los argumentos hechos por aquellos que difieren con la conclusión del IPCC, y evalúe cada argumento.
3. Revise los niveles legales de confianza en causalidad estudiada en la sección 7.4. ¿Diría que el caso para actividad humana como la causa de calentamiento global está ahora en el nivel de causa posible, causa probable o causa más allá de cualquier duda razonable? Defienda su opinión.
4. Con base en su punto de vista del calentamiento global describa lo que considere que debe hacerse con él.

LECTURAS RECOMENDADAS

Bennet, Jeffrey y Shostak, Seth, *Life in the Universe*, segunda edición, Addison Wesley, 2007 (vea la sección 10.5).

Flannery, Tim, *The Weather Makers*, Atlantic Monthly Press, 2005.

IPCC, *Climate Change 2007*, El reporte completo está disponible en línea en <http://www.ipcc.ch>.



Algunas personas odian todo lo que suene a estadísticas, pero yo lo encuentro completamente bello e interesante. Siempre que no sean embrutecidas, sino manejadas con delicadeza por métodos superiores, y sean interpretadas con cautela, su potencia de tratar con fenómenos complicados es extraordinaria.

—Sir Francis Galton (1822-1912)

De muestras a poblaciones

¿ALGUNA VEZ SE HA MARAVILLADO DE CÓMO EL RESULTADO DE UNA ELECCIÓN NACIONAL PUEDE PREDECIRSE HORAS ANTES DE QUE SE CIERREN LAS CASILLAS? O, ¿CÓMO UN GRAN VENDEDOR DE BIENES AL MENUDEO PUEDE TOMAR DECISIONES CRUCIALES CON BASE EN UNA ENCUESTA A SÓLO UNOS CUANTOS CIENTOS DE PERSONAS? QUIZÁS ESTOS EJEMPLOS ILUSTRAN EL ASPECTO MÁS PODEROSO DE LA ESTADÍSTICA: LA CAPACIDAD DE UTILIZAR INFORMACIÓN REUNIDA DE UNA MUESTRA PEQUEÑA PARA HACER PREDICCIONES ACERCA DE UNA POBLACIÓN MUCHO MÁS GRANDE. ESTE PROCESO, DENOMINADO *CONSTRUCCIÓN DE INFERENCIAS*, ES EL TEMA DE LA RAMA DE LA ESTADÍSTICA LLAMADA *INFERENCIA ESTADÍSTICA*.

¿ALGUNA VEZ SE HA MARAVILLADO DE CÓMO EL RESULTADO DE UNA ELECCIÓN NACIONAL PUEDE PREDECIRSE HORAS ANTES DE QUE SE CIERREN LAS CASILLAS? O, ¿CÓMO UN GRAN VENDEDOR DE BIENES AL MENUDEO PUEDE TOMAR DECISIONES CRUCIALES CON BASE EN UNA ENCUESTA A SÓLO UNOS CUANTOS CIENTOS DE PERSONAS? QUIZÁS ESTOS EJEMPLOS ILUSTRAN EL ASPECTO MÁS PODEROSO DE LA ESTADÍSTICA: LA CAPACIDAD DE UTILIZAR INFORMACIÓN REUNIDA DE UNA MUESTRA PEQUEÑA PARA HACER PREDICCIONES ACERCA DE UNA POBLACIÓN MUCHO MÁS GRANDE. ESTE PROCESO, DENOMINADO *CONSTRUCCIÓN DE INFERENCIAS*, ES EL TEMA DE LA RAMA DE LA ESTADÍSTICA LLAMADA *INFERENCIA ESTADÍSTICA*.

OBJETIVOS DE APRENDIZAJE

8.1 Distribuciones de muestreo

Entender las ideas fundamentales de las distribuciones de muestreo y cómo la distribución de las medias y la distribución de las proporciones muestrales están formadas. También aprender la notación usada para representar medias muestrales y proporciones.

8.2 Estimación de medias poblacionales

Aprender a estimar medias poblacionales y calcular los márgenes de error e intervalos de confianza.

8.3 Estimación de proporciones poblacionales

Aprender a estimar las proporciones poblacionales y a calcular los márgenes asociados de error en intervalos de confianza.

8.1 Distribuciones de muestreo

Considere los enunciados siguientes tomados de artículos recientes o de reportes de investigación.

- El consumo promedio diario de proteína por los estadounidenses es de 67 gramos.
- En todo el país, la estancia media en el hospital después del parto de un bebé disminuyó de 3.2 días en 1980 a la media actual de 2.0 días.
- Treinta por ciento de las alumnas de preparatoria en Estados Unidos creen que serían más felices casadas que no siendo casadas.
- Alrededor de 5% de todos los niños estadounidenses viven con un abuelo.

Todos los días hemos oído o leído enunciados como éstos. Ellos hacen una afirmación acerca de una población grande de individuos, aunque no sea posible que todos en la población hayan sido encuestados o medidos. Por ejemplo, el tercer enunciado estuvo basado en una muestra aleatoria de alumnas de preparatoria, no estuvo basado en todas las alumnas de preparatoria del país. Las respuestas de las alumnas en la muestra fueron usadas para hacer una afirmación acerca de toda la población de alumnas de preparatoria. ¿Cómo es posible *inferir* de una muestra aleatoria una conclusión muy general acerca de la población? Estas preguntas van al corazón de la rama de la estadística denominada *estadística inferencial*.

Observe que en los enunciados anteriores se hicieron dos tipos de afirmaciones. Los primeros dos enunciados dan estimaciones de la *media* de una cantidad —consumo medio de proteínas de 67 gramos y días promedio de estancia de 3.2 y 2.0 días—. Los últimos dos enunciados dicen algo acerca de una *proporción* de la población, 30% de alumnas de preparatoria y 5% de niños estadounidenses. Estas medias y proporciones pertenecen a toda la población, por lo que son *parámetros de la población* (vea la sección 1.1). El tema de este capítulo, y uno de los objetivos principales de la estadística inferencial, es aprender cómo estimar los parámetros de la población usando datos de muestras.

En este país la opinión pública lo es todo.

—Abraham Lincoln

Medias muestrales: la idea básica

Gran parte del trabajo en éste y en el capítulo siguiente incluye la selección de una muestra de una población, analizar la muestra y luego sacar conclusiones acerca de la población con base en lo que se aprendió de la muestra. Para ganar alguna experiencia en este importante proceso, iniciaremos con un ejemplo.

La tabla 8.1 lista los pesos de los cinco jugadores titulares (por conveniencia rotulados de la A a la E) en un equipo profesional de básquetbol. Para los propósitos de este ejemplo, consideramos estos cinco jugadores como la *población completa* (con una media de 242.4 libras); en otras palabras, es la población completa de titulares del equipo. Ésta es una población muy pequeña para los estándares estadísticos, pero su tamaño pequeño nos permitirá observar cuidadosamente el proceso de muestreo. El tamaño de las muestras sacadas de esta población de cinco jugadores van de $n = 1$ (un jugador de los cinco) a $n = 5$ (los cinco jugadores).

Con un tamaño de muestra de $n = 1$, existen 5 diferentes muestras que podríamos seleccionar: cada jugador es una muestra. La media de cada muestra de tamaño $n = 1$ es simplemente el peso del jugador en la muestra. La figura 8.1 muestra un histograma de las medias de las 5 muestras; se denomina **distribución de las medias muestrales**, ya que muestra las medias de todas las muestras de tamaño $n = 1$. La distribución de las medias muestrales mediante este proceso es un ejemplo de una **distribución muestral**. Este término se refiere simplemente a una distribución de una estadística muestral, tal como la media, tomada de todas las muestras

posibles de un tamaño particular. Observe que la media de las 5 medias muestrales es la media de toda la población:

$$\frac{215 + 242 + 225 + 215 + 315}{5} = 242.4 \text{ libras}$$

Esto demuestra una regla general: *la media de una distribución de medias muestrales es la media de la población.*

Tabla 8.1 Pesos para la población de cinco titulares en un equipo de básquetbol	
Jugador	Peso (libras)
A	215
B	242
C	225
D	215
E	315

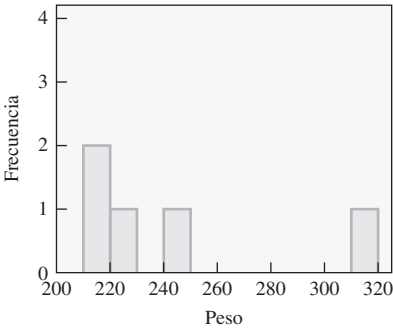


Figura 8.1 Distribución de muestreo para tamaño de muestra $n = 1$.

Pasemos a muestras de tamaño $n = 2$, en las que cada muestra consiste en dos jugadores diferentes. Con cinco jugadores hay 10 muestras diferentes de tamaño $n = 2$. Cada muestra tiene su propia media. Por ejemplo, la muestra de jugadores A y B tiene una media de $(215 + 242)/2 = 228.5$ libras. La tabla 8.2 lista las 10 muestras con sus medias y la figura 8.2 muestra la distribución de las 10 medias muestrales. Nuevamente, observe que la media de la distribución de las medias muestrales es igual a la media de la población, 242.4 libras.

Tabla 8.2 Medias muestrales para el ejemplo del básquetbol; tamaño de la muestra $n = 2$	
Muestra	Media
AB	228.5
AC	220.0
AD	215.0
AE	265.0
BC	233.5
BD	228.5
BE	278.5
CD	220.0
CE	270.0
DE	265.0
Media	242.4

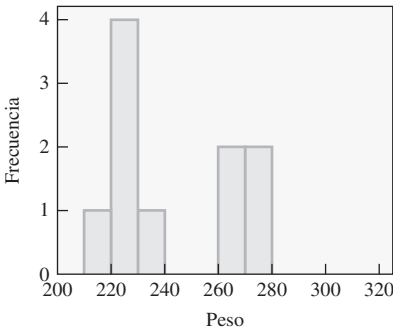


Figura 8.2 Distribución de muestreo para tamaño de muestra $n = 2$.

NOTA TÉCNICA

Las tablas 8.2, 8.3 y 8.4 incluyen muestreo *sin reemplazo*, que tiene la ventaja de evitar la duplicidad que ocurre cuando el mismo elemento es seleccionado más de una vez. Sin embargo, en estadística estamos interesados particularmente en la selección *con reemplazo* por las razones siguientes: en poblaciones grandes no hay diferencia significativa si el muestreo es con o sin reemplazo, pero el muestreo con reemplazo da como resultado eventos independientes que no son afectados por resultados anteriores; los eventos independientes son más sencillos de analizar y dan lugar a fórmulas sencillas.

Son posibles diez muestras de tamaño $n = 3$ en una población de cinco jugadores. La tabla 8.3 presenta estas muestras y sus medias, y la figura 8.3 da la distribución de estas medias muestrales. De nuevo, la media de la distribución de las medias muestrales es igual a la media de la población, 242.4 libras.

Tabla 8.3 Medias muestrales para el ejemplo del básquetbol; tamaño de la muestra $n = 3$	
Muestra	Media
ABC	227.3
ABD	224.0
ABE	257.3
ACD	218.3
ACE	251.7
ADE	248.3
BCD	227.3
BCE	260.7
BDE	257.3
CDE	251.7
Media	242.4

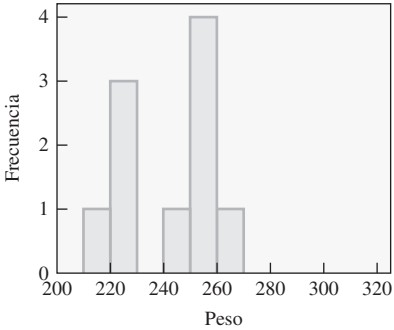


Figura 8.3 Distribución de muestreo para tamaño de muestra $n = 3$.

Con un tamaño de muestra de $n = 4$, sólo son posibles 5 muestras diferentes. La tabla 8.4 presenta estas muestras y sus medias, y la figura 8.4 da la distribución de estas medias muestrales.

Tabla 8.4 Medias muestrales para el ejemplo del básquetbol; tamaño de la muestra $n = 4$	
Muestra	Media
ABCD	224.25
ABCE	249.25
ABDE	246.75
ACDE	242.50
BCDE	249.25
Media	242.4

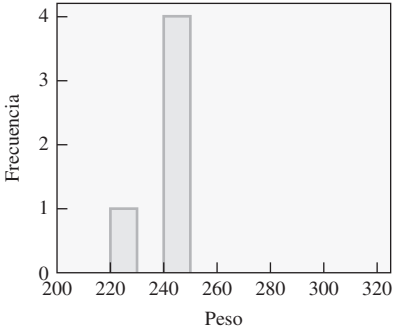


Figura 8.4 Distribución de muestreo para tamaño de muestra $n = 4$.

Por último, para una población de cinco jugadores, sólo existe 1 muestra posible de tamaño $n = 5$, la población completa. En este caso, la distribución de las medias muestrales es una sola barra (figura 8.5). Nuevamente la media de la distribución de las medias muestrales es la media de la población, 242.4 libras.

Para resumir, cuando trabajamos con *todas* las muestras posibles de una población de un tamaño dado, la media de la distribución de medias muestrales es siempre la media de la población. Una mirada más cuidadosa a las distribuciones en las figura 8.1 a 8.5 también revela que conforme el tamaño de la muestra aumenta, la distribución se hace más angosta y se agrupa alrededor de la media. De hecho, si observamos en una población grande y tamaños de muestra grandes, encontraríamos la distribución de las medias muestrales cada vez más parecida a una distribución normal, conforme el tamaño de la muestra aumenta (una consecuencia del teorema del límite central, analizado en la sección 5.3).

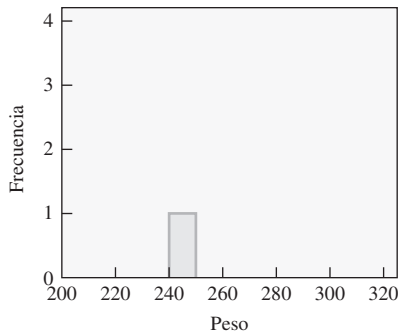


Figura 8.5 Distribución de muestreo para tamaño de muestra $n = 5$.

Medias muestrales con poblaciones grandes

En estudios estadísticos reales, es muy raro que trabajemos con poblaciones tan pequeñas de cinco individuos. Para explorar el proceso de selección de muestras y la formación de distribuciones de medias muestrales en un ambiente más realista, consideramos una población un poco más grande.

Imagine que trabaja para el departamento de ciencias de la computación de una pequeña universidad. Para desarrollar estrategias de conexión de redes, encuesta a 400 estudiantes de la universidad para determinar cuántas horas a la semana destinan a realizar búsquedas en internet. Las respuestas (en horas por semana) se muestran a continuación:

3.4	6.8	6.7	3.4	0.0	5.0	5.4	1.8	0.7	1.6	2.1	3.5	3.4	6.4	7.2	1.8	7.4	3.0	4.0	5.2
1.2	7.8	7.0	0.4	7.2	4.8	3.6	8.0	5.4	6.4	3.5	5.3	4.7	5.4	5.6	3.8	0.1	2.4	0.5	4.0
4.5	8.0	4.2	1.0	6.2	7.1	3.8	0.7	5.5	1.7	2.6	1.6	0.7	1.3	6.5	2.4	3.0	0.3	2.2	0.4
1.9	5.0	2.0	5.3	7.5	5.0	0.3	7.4	6.0	4.3	1.3	0.8	7.2	6.6	0.2	3.4	1.6	2.2	3.0	4.5
5.5	5.3	6.5	0.1	0.3	4.2	2.2	6.2	7.3	3.1	5.4	1.3	6.3	4.5	7.1	5.8	6.1	0.5	0.4	4.1
7.0	6.0	1.1	0.8	1.4	2.9	7.3	0.8	2.7	0.6	3.0	0.7	2.8	6.5	1.9	3.6	1.6	2.6	2.6	6.6
6.8	6.1	3.6	1.4	7.7	5.2	3.8	6.0	2.2	7.5	6.7	4.4	4.1	7.3	5.2	5.7	6.7	2.4	0.6	6.7
1.0	2.3	0.7	1.2	4.5	3.3	4.2	2.1	5.9	3.0	7.2	7.9	2.5	7.1	8.0	6.7	4.1	4.9	0.0	3.1
6.0	0.5	4.2	2.7	0.1	1.4	2.1	2.5	3.9	5.8	5.9	2.7	2.8	3.7	7.3	0.7	6.9	4.4	0.7	1.6
3.1	2.1	7.4	3.6	6.5	2.9	5.4	3.9	3.0	0.8	0.3	0.8	3.3	0.8	8.0	5.6	7.1	1.3	0.2	5.2
7.8	4.7	7.2	0.9	5.1	0.9	1.7	1.2	0.4	6.9	0.6	3.0	3.6	6.1	1.6	6.0	3.8	0.4	1.1	4.0
3.8	4.0	1.8	0.9	1.1	3.9	1.7	1.7	2.6	0.1	4.0	1.4	1.9	0.9	0.2	4.2	4.7	0.2	5.3	2.2
5.8	7.5	5.8	5.2	3.9	3.4	7.3	4.1	0.5	7.9	7.7	7.7	5.0	2.3	7.8	2.3	5.6	6.5	7.9	5.0
2.0	5.5	5.4	6.6	6.7	4.4	7.2	2.5	4.9	7.0	2.1	7.2	4.1	1.2	6.2	3.3	6.3	2.3	4.9	2.2
6.4	7.2	0.1	5.3	3.0	0.7	1.5	1.2	1.1	7.4	5.1	7.2	7.2	3.0	7.1	4.5	6.7	7.2	7.2	0.9
2.9	4.3	2.5	0.7	7.6	3.9	0.7	5.8	6.6	3.4	0.3	6.5	7.5	0.7	6.1	6.1	4.8	1.9	1.9	5.0
1.1	7.8	6.8	4.9	3.0	6.5	5.2	2.2	5.1	3.4	4.7	7.0	3.8	5.7	6.8	1.2	1.7	6.5	0.1	4.3
6.3	1.2	0.8	0.7	0.6	7.0	4.0	6.6	6.9	0.5	4.3	1.0	0.5	3.1	0.9	2.3	5.7	6.7	7.3	0.5
0.3	0.9	2.4	2.5	7.8	5.6	3.2	0.7	5.4	0.0	5.7	0.3	7.2	5.1	2.5	3.2	3.1	2.8	5.0	5.6
3.1	0.7	0.5	3.9	2.6	7.3	1.4	1.2	7.1	5.5	3.1	5.0	6.8	6.5	1.7	2.1	7.3	4.0	2.2	5.6

Por supuesto que podría calcular la media de las 400 respuestas. Encontraría que es 3.88 horas a la semana. Esta media es la **media poblacional** verdadera, ya que es la media de toda la población de 400 estudiantes; denotamos a la media de la población con la letra griega μ (se pronuncia “mu”). De forma similar, un cálculo muestra que la desviación estándar de la población es $\sigma = 2.40$.

En aplicaciones comunes de estadística, las poblaciones son enormes y es poco práctico o costoso encuestar a cada individuo de la población; en consecuencia, es raro que conozcamos la media poblacional verdadera, μ . Por tanto, tiene sentido considerar el uso de la media de una *muestra* para estimar la media de toda la población. Aunque es más sencillo de trabajar con

una muestra, es posible que no represente de manera exacta a toda la población. De modo que no debemos esperar que una estimación de la media de la población obtenida a través de una muestra sea perfecta. El error que introducimos al trabajar con una muestra se denomina **error de muestreo**. Podemos explorar esta idea considerando muestras sacadas de las 400 respuestas acerca del uso de internet.

Error de muestreo

El **error de muestreo** es el error introducido a consecuencia de que una muestra aleatoria es usada para estimar un parámetro de la población. No incluye otras fuentes de error, tales como las debidas a sesgo en el muestreo, malas preguntas en las encuestas o errores registrados.

UN MOMENTO DE REFLEXIÓN

¿Esperaría que el error de muestreo aumentase o disminuyese si el tamaño de la muestra estuviera aumentando? Explique.

Suponga que selecciona una muestra aleatoria de $n = 32$ respuestas del conjunto de datos de la página 337 y calcula su media. Por ejemplo, la muestra aleatoria podría ser

1.1 7.8 6.8 4.9 3.0 6.5 5.2 2.2 5.1 3.4 4.7 7.0 3.8 5.7 6.5 2.7
2.6 1.4 7.1 5.5 3.1 5.0 6.8 6.5 1.7 2.1 1.2 0.3 0.9 2.4 2.5 7.8

La media de esta muestra es $\bar{x} = 4.17$; utilizamos la notación estándar $\bar{x}$ para denotar a esta media. Decimos que $\bar{x}$ es un *estadístico muestral* ya que proviene de una muestra de la población entera. Así, $\bar{x}$ se denomina **media muestral**.

Notación para las medias poblacional y muestral

n = tamaño de la muestra

μ = media poblacional

$\bar{x}$ = media muestral

Nuestro objetivo es utilizar una sola media muestral para estimar la media poblacional, pero primero es útil analizar otras muestras. Suponga que recolecta muestras adicionales con el propósito de estudiar el comportamiento de las medias muestrales. Aquí está un ejemplo:

1.8 0.4 4.0 2.4 0.8 6.2 0.8 6.6 5.7 7.9 2.5 3.6 5.2 5.7 6.5 1.2
5.4 5.7 7.2 5.1 3.2 3.1 5.0 3.1 0.5 3.9 3.1 5.8 2.9 7.2 0.9 4.0

Para este ejemplo, la media muestral es $\bar{x} = 3.98$.

Ahora tenemos dos medias muestrales que no coinciden una con la otra, y ninguna coincide con la media poblacional verdadera. Suponga que decide seleccionar muchas más muestras de 32 respuestas (una opción que por lo regular no existe en la práctica). Para cada muestra usted calcula la media muestral, $\bar{x}$. Con el fin de obtener una buena idea de todas estas medias muestrales, podría construir un histograma para indicar el número de muestras con medias muestrales distintas. La figura 8.6 muestra un histograma que resulta de 100 muestras diferentes, cada una con 32 estudiantes. Observe que este histograma es muy parecido a una *distribución normal* y su media es muy parecida a la media de la población, $\mu = 3.88$.

UN MOMENTO DE REFLEXIÓN

Suponga que sólo seleccionó una muestra de tamaño $n = 32$. De acuerdo con la figura 8.6, ¿es más probable elegir una muestra con media menor que 2.5 o una muestra con media menor que 3.5? Explique.

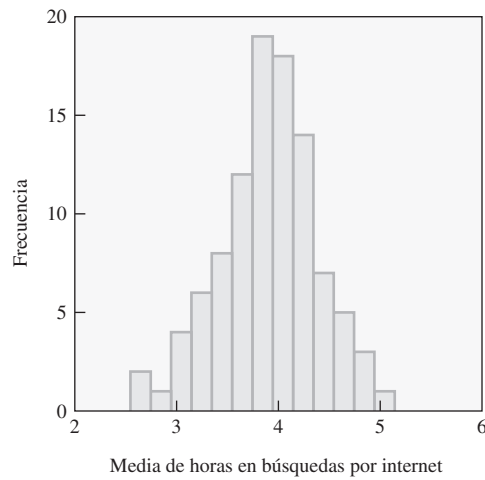


Figura 8.6 Una distribución de 100 medias muestrales, con un tamaño de muestra $n = 32$, parece muy cercana a una distribución normal con una media de 3.88.

En nuestro ejemplo inicial con la población de cinco jugadores de básquetbol, fuimos capaces de examinar *todas* las muestras posibles de cada uno de los tamaños desde $n = 1$ hasta $n = 5$. No podemos realizar la misma tarea para una población de 400, ya que el número de posibles muestras (combinaciones) de tamaño $n = 32$ es cercano 2×10^{47} . Sin embargo, incluso con el número pequeño de muestras de la figura 8.6, ya podemos ver que la distribución de las medias muestrales es casi normal. Una vez más, estamos viendo una propiedad importante de la distribución de las medias muestrales. La distribución de las medias muestrales se aproxima a una distribución normal para tamaños grandes de muestras y la media para la distribución de las medias muestrales es igual a la media de la población, μ .

En la práctica, con frecuencia sólo se dispone de *una* muestra. En ese caso, su media muestral, $\bar{x}$, es nuestra mejor estimación de la media poblacional, μ . Por fortuna, como veremos en breve, todavía es posible decir algo acerca de cuán bien $\bar{x}$ aproxima a μ .

La distribución de las medias muestrales

La **distribución de las medias muestrales** es la distribución que resulta cuando encontramos las medias de *todas* las posibles muestras de un tamaño dado. Entre mayor sea el tamaño de la muestra, esta distribución se aproxima más a una distribución normal. En todos los casos, la media de la distribución de las medias muestrales es igual a la media de la población. Si sólo está disponible una muestra, su media muestral, $\bar{x}$, es la mejor estimación para la media de la población, μ .

NOTA TÉCNICA

Una guía común es suponer que la distribución de las medias muestrales es casi normal si el tamaño de la muestra es mayor que 30.

Podemos ampliar esta idea para ver el importante papel que desempeña la distribución normal. De nuevo considere la distribución de las medias muestrales de la figura 8.6. Si estuvieran incluidas *todas* las muestras posibles de tamaño $n = 32$, esta distribución tendría estas características:

- La distribución de las medias muestrales es aproximadamente una distribución normal.
- La media de la distribución de las medias muestrales es 3.88 (la media de la población).
- La desviación estándar de la distribución de las medias muestrales depende de la desviación estándar de la población y del tamaño de la muestra. La desviación estándar de la población es $\sigma = 2.40$ y el tamaño de la muestra es $n = 32$, por lo que la desviación estándar de las medias muestrales es

$$\frac{\sigma}{\sqrt{n}} = \frac{2.40}{\sqrt{32}} = 0.42$$

Suponga que seleccionamos la siguiente muestra aleatoria de 32 respuestas de las 400 respuestas dadas anteriormente:

5.8 7.5 5.8 5.2 3.9 3.4 7.3 4.1 0.5 7.9 7.7 7.7 5.0 2.3 7.8 2.3
5.0 6.8 6.5 1.7 2.1 7.3 4.0 2.2 5.6 4.7 5.3 3.5 6.5 3.4 6.6 5.0

La media de esta muestra es $\bar{x} = 5.01$. Dado que la media de la distribución de medias muestrales es 3.88 y la desviación estándar es 0.42, la media muestral de $\bar{x} = 5.01$ tiene una *puntuación estándar* de

$$z = \frac{\text{media muestral} - \text{media pob.}}{\text{desviación estándar}} = \frac{5.01 - 3.88}{0.42} = 2.7$$

(Recuerde que utilizamos z para denotar una puntuación estándar; vea la sección 5.2 para revisión). La muestra que hemos seleccionado tiene una puntuación estándar de $z = 2.7$, indicando que está 2.7 desviaciones estándar por arriba de la media de la distribución muestral. En la tabla 5.1 esta puntuación estándar corresponde al percentil número 99.65, por lo que la probabilidad de seleccionar otra muestra con una media *menor* que 5.01 es alrededor de 0.9965. Se sigue que la probabilidad de seleccionar otra muestra con una media *mayor* que 5.01 es alrededor de $1 - 0.9965 = 0.0035$. Aparentemente, la muestra que seleccionamos es muy extrema dentro de esta población. Este ejemplo muestra que cuando tenemos una distribución de muestreo normal, podemos determinar si una muestra particular es un tanto ordinaria o es muy rara.

A propósito...

El número total de granjas en Estados Unidos ha descendido de 2 440 000 en 1980 a 2 172 000 en 2000. La media del número de acres por granja aumentó de 426 en 1980 a 434 en 2000.



UN MOMENTO DE REFLEXIÓN

Suponga que una media muestral está en el percentil 95o. Explique por qué la probabilidad de seleccionar de manera aleatoria otra muestra con una media mayor que la primera media es 0.05.

EJEMPLO 1 Muestreo de granjas

Texas tiene aproximadamente 225 000 granjas, más que cualquier otro estado en Estados Unidos. El tamaño medio real de las granjas es $\mu = 582$ acres y la desviación estándar es $\sigma = 150$ acres. Para muestras aleatorias de $n = 100$ granjas, se determinó la media y la desviación estándar de la distribución de las medias muestrales. ¿Cuál es la probabilidad de seleccionar una muestra aleatoria de 100 granjas con una media mayor a 600 acres?

Solución Puesto que la distribución de las medias muestrales es una distribución normal, su media debe ser la misma que la media de toda la población, que es 582 acres. La desviación estándar de la distribución muestral es $\sigma/\sqrt{n} = 150/\sqrt{100} = 15$. Por tanto, una media muestral de $\bar{x} = 600$ acres tiene una puntuación estándar de

$$z = \frac{\text{media muestral} - \text{media pob.}}{\text{desviación estándar}} = \frac{600 - 582}{15} = 1.2$$

De acuerdo con la tabla 5.1, esta puntuación estándar está en el percentil 88o., así que la probabilidad de seleccionar una muestra con una media menor que 600 acres es alrededor de 0.88. Por tanto, la probabilidad de seleccionar una muestra con una media mayor que 600 acres es alrededor de 0.12.

Proporciones muestrales

Mucho de lo que hemos aprendido acerca de las distribuciones de medias muestrales lleva a la distribución de *proporciones muestrales*. Podemos ver el paralelismo regresando a la encuesta de los 400 estudiantes, que se describió antes. Suponga que su objetivo es determinar la proporción (o porcentaje) de los 400 estudiantes que son propietarios de un automóvil. A continuación, cada Y (por *sí*) o N (por *no*) es la respuesta de una persona a la pregunta “¿Usted es propietario de un automóvil?”.

Y N Y Y N Y N N Y N Y Y N N N Y Y Y Y N Y N Y N Y Y Y N Y Y N Y Y N N N Y Y Y N
 Y N Y N Y N Y Y Y Y Y N Y Y N Y Y N N N Y Y Y N Y N Y N Y Y Y Y N Y Y N
 Y N N Y N Y Y N N N Y Y Y Y N Y N Y N Y Y Y Y N Y N Y N Y Y Y N Y Y N Y Y N
 N N Y Y Y N Y N Y N Y N Y Y Y Y N Y Y N Y N N Y N Y N N N N N Y Y Y N Y N Y
 N Y N Y Y Y Y N Y Y N Y N N Y N Y Y N N N Y Y Y N Y N Y N Y Y Y N Y Y N Y Y
 N Y Y N Y N N Y N Y Y N N N Y Y Y N N N Y Y Y N Y N Y N Y Y Y Y N Y N Y
 N Y Y Y N Y N Y Y N Y Y Y N Y Y N Y Y Y N N N Y Y Y N N N Y Y Y N Y N Y N Y
 Y N Y N Y N N Y N Y N Y N Y Y Y Y Y N Y N Y N Y Y Y Y N Y N Y N Y Y N Y N
 N Y N Y Y N N N Y Y Y N N N Y Y Y N Y N N Y Y Y N Y Y Y Y Y N Y N Y N Y Y
 Y N Y Y N Y N N Y N Y Y N N N Y Y Y N N N Y Y Y N Y N Y N Y Y Y N Y Y N Y Y

Si cuenta con cuidado, encontrará que 240 de las 400 respuestas son Y, por lo que la proporción *exacta* de propietarios de automóviles en la población de 400 estudiantes es

$$p = \frac{240}{400} = 0.6$$

Esta *proporción poblacional*, $p = 0.6$, es otro ejemplo de un *parámetro poblacional*. De los 400 estudiantes en la población, ésta es la proporción verdadera de propietarios de automóviles.

UN MOMENTO DE REFLEXIÓN

Proporcione otra pregunta de encuesta que resultaría en una proporción poblacional en lugar de una media poblacional.

Una vez más, en problemas comunes de estadística, es poco práctico o muy caro encuestar a todos los individuos de una población. Por tanto, es razonable considerar la idea de utilizar una muestra aleatoria de, digamos, $n = 32$ personas para estimar la proporción poblacional. Suponga que usted de manera aleatoria saca 32 respuestas de la lista de Y y N para generar la muestra siguiente:

Y N Y Y N Y Y N Y N Y Y N Y Y Y N N Y Y Y N N Y Y N Y Y Y N Y Y

La proporción de respuestas Y en esta lista es

$$\hat{p} = \frac{21}{32} = 0.656$$

Esta proporción es otro ejemplo de *estadística muestral*. En este caso, es una **proporción muestral** ya que es la proporción de propietarios de automóvil dentro de una *muestra*; utilizamos el símbolo $\hat{p}$ (se lee “ p gorro”) para distinguir esta proporción muestral de la proporción poblacional, p .

Notación para las proporciones de la población y de la muestra

n = tamaño de la muestra

p = proporción de la población

$\hat{p}$ = proporción muestral

Nuestro objetivo es determinar cuán bien una *sola* proporción muestral, $\hat{p}$, aproxima la proporción de la población, p . Por el momento, imagine que se puede dar el lujo de seleccionar otra muestra de $n = 32$ respuestas; esta muestra produce la lista

Y Y N N N Y Y N Y Y Y N N Y N Y Y N Y N Y Y N Y Y N N Y N Y

La proporción de respuestas Y en la lista es

$$\hat{p} = \frac{18}{32} = 0.563$$

Existen tres clases de opiniones: las informadas, las desinformadas y las intrascendentes. Las opiniones desinformadas abarcan la mayoría de las opiniones de estadounidenses.

—Paul Talmey, encuestador

Suponga que decide seleccionar *muchas* más muestras de 32 respuestas (otra vez, una opción que por lo común no ocurre en la práctica). Para cada muestra, usted calcula la proporción muestral, $\hat{p}$. Si dibuja un histograma de muchas proporciones muestrales, mostrará el número de muestras que tienen valores particulares de $\hat{p}$ de 0 a 1. La figura 8.7 es uno de tales histogramas, en este caso el resultado de 100 muestras de tamaño $n = 32$. Como encontramos para las medias muestrales, esta distribución de proporciones muestrales es muy cercana a una distribución normal. Además, la media de esta distribución es muy cercana a la proporción poblacional de 0.6.

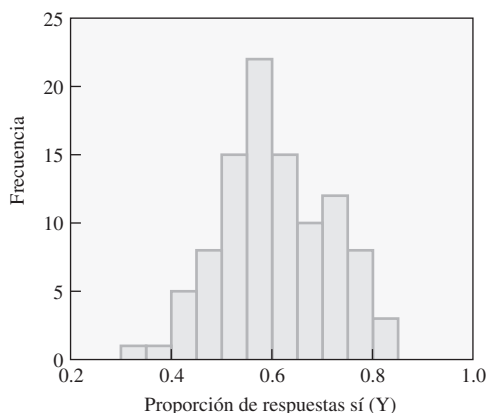


Figura 8.7 La distribución de 100 proporciones muestrales, con un tamaño de muestra de 32, parece que es cercana a la distribución normal.

Suponga que fuese posible seleccionar *todas* las posibles muestras de tamaño $n = 32$. La distribución resultante sería llamada **distribución de proporciones muestrales**. La media de esta distribución es igual a la proporción exacta de la población, y esta distribución se aproxima a una distribución normal cuando el tamaño de la muestra crece.

En la práctica, con frecuencia sólo tenemos una muestra para trabajar con ella. En ese caso, la mejor estimación para la proporción poblacional, p , es la proporción muestral, $\hat{p}$.

La distribución de las proporciones muestrales

La **distribución de las proporciones muestrales** es la distribución que resulta cuando encontramos las proporciones $\hat{p}$ en *todas* las posibles muestras de un tamaño dado. Entre mayor sea el tamaño de la muestra, esta distribución se aproxima más a la distribución normal. En todos los casos, la media de la distribución de las proporciones muestrales es igual a la proporción de la población. Si sólo está disponible una muestra, su proporción muestral, $\hat{p}$, es la mejor estimación para la proporción poblacional, p .

EJEMPLO 2 Análisis de una proporción muestral

Considere la distribución de las proporciones muestrales que aparece en la figura 8.7. Suponga que su media es $p = 0.6$ y su desviación estándar es 0.1. Suponga que selecciona de manera aleatoria la muestra siguiente de 32 respuestas:

Y Y N Y Y Y Y N Y Y Y Y Y N Y Y N Y Y Y N Y Y N Y Y N Y N Y Y

Calcule la proporción muestral, $\hat{p}$, para esta muestra. ¿Qué tan lejos está de la media de la distribución? ¿Cuál es la probabilidad de seleccionar otra muestra con una proporción mayor que la que seleccionó?

Solución En este ejemplo, la proporción de respuestas Y es

$$\hat{p} = \frac{24}{32} = 0.75$$

Usando la media de 0.6 y una desviación estándar de 0.1, encontramos que el estadístico muestral, $\hat{p} = 0.75$, tiene una puntuación estándar de

$$z = \frac{\text{proporción muestral} - \text{proporción pob.}}{\text{desviación estándar}} = \frac{0.75 - 0.6}{0.1} = 1.5$$

La proporción muestral está 1.5 desviaciones estándar por arriba de la media de la distribución. Usando la tabla 5.1 vemos que una puntuación estándar de 1.5 corresponde al percentil 93o. La probabilidad de seleccionar otra muestra con una proporción menor que la seleccionada es alrededor de 0.93. Así, la probabilidad de seleccionar otra muestra con una proporción mayor a la que seleccionamos es alrededor de $1 - 0.93 = 0.07$. En otras palabras, si seleccionáramos 100 muestras aleatorias de 32 respuestas, esperaríamos ver sólo 7 muestras con una proporción mayor que la que seleccionamos.

NOTA TÉCNICA

La desviación estándar de proporciones muestrales es

$$\sqrt{\frac{p(1-p)}{n}}$$

Si no conocemos el valor de la proporción de la población, p , podemos estimar la desviación estándar usando la proporción muestral, $\hat{p}$. Por ejemplo, $\hat{p} = 0.75$, si $n = 32$, obtenemos

$$\sqrt{\frac{0.75(1 - 0.75)}{32}} = 0.0765$$

o redondeado 0.1.

Sección 8.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Distribución muestral.** Con frecuencia los encuestadores utilizan dígitos, entre 0 y 9, seleccionados de manera aleatoria para generar las partes de números telefónicos para llamar. ¿Cuál es la distribución de tales dígitos seleccionados aleatoriamente? Si repetimos el proceso de generar aleatoriamente 50 dígitos y determinar la media, ¿cuál es la distribución resultante de las medias muestrales?
- Mejores estimaciones.** Dada una sola muestra, ¿cuál es la mejor estimación de la media poblacional? Dada una sola muestra, ¿cuál es la mejor estimación de la proporción muestral?
- Notación.** ¿Qué denota $\bar{x}$?, ¿qué denota μ ? y ¿cuál es la diferencia entre ellas?
- Notación.** ¿Qué denota $\hat{p}$?, ¿qué denota p ? y ¿cuál es la diferencia entre ellas?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Muestra grande.** Ted es un estudiante universitario que trabaja en un proyecto para estimar la proporción de estadounidenses adultos que han estado involucrados en un accidente automovilístico. Utiliza una muestra de conveniencia de estudiantes de su escuela, pero obtiene una muestra muy grande de tamaño 3000. Ted establece que su estimación

es buena ya que el gran tamaño de la muestra compensa el hecho que está usando muestreo de conveniencia.

- Errores registrados.** Un error de muestreo ocurre cuando un encuestador, de manera incorrecta, registra una respuesta de una encuesta.
- Tamaño grande de muestra.** Cuando un muestreo aleatorio es utilizado para estimar una media poblacional, la media muestral tiende a volverse una mejor estimación de la media poblacional conforme el tamaño de la muestra aumenta.
- Encuesta.** En una encuesta de Gallup a 1012 adultos elegidos de manera aleatoria, 901 dijeron que debe permitirse la clonación de humanos. Si fuese obtenida una muestra grande, la proporción muestral sería probablemente mayor que 901/1012 (o 0.890).

Conceptos y aplicaciones

- Estimación de proporciones de una población.** En una encuesta de niños de 5 a 17 años, 1050 fueron seleccionados aleatoriamente de los nueve estados del noreste, y la proporción que en su casa hablaba un idioma distinto del inglés fue 0.19. Con base en este estadístico muestral, ¿cuál es la mejor estimación de la proporción de todos los niños en el noreste? ¿La proporción muestral es una buena estimación para los niños de Estados Unidos? ¿Por qué sí o por qué no?
- Estimación de medias de una población.** Cuando 40 mujeres fueron seleccionadas de manera aleatoria y se les hizo un examen de niveles de colesterol; se obtuvo una media

de 240.9 miligramos. Con base en este estadístico muestral, ¿cuál es la mejor estimación para la media del nivel de colesterol en todas las mujeres? ¿Estaría más confiado de su estimación si la muestra incluyese las mediciones de 500 mujeres? Explique.

- 11. Distribución de medias muestrales.** Suponga que se llenan latas de coca de modo que las cantidades reales tienen una media de 12.00 onzas. Una muestra aleatoria de 36 latas tiene una cantidad media de 12.19 onzas. La distribución de las medias muestrales de tamaño 36 es normal con una media supuesta de 12.00 onzas y una desviación estándar de 0.02 onzas.

- ¿A cuántas desviaciones estándar está la media muestral de la media de la distribución de medias muestrales?
- En general, ¿cuál es la probabilidad de que una muestra aleatoria de tamaño 36 tenga una media de al menos 12.19 onzas?
- ¿Parece que los consumidores son engañados? ¿Por qué sí o por qué no?

- 12. Distribución de medias muestrales.** Suponga que sabe que la distribución de las medias muestrales para tamaños de familias para muestras de 500 hogares es normal con una media de 2.64 y una desviación estándar de 0.06. Suponga que selecciona una muestra aleatoria de $n = 500$ hogares y determina que el número medio de personas por hogar, para esta muestra, es 2.55.

- ¿A cuántas desviaciones estándar está la media muestral de la media de la distribución de medias muestrales?
- ¿Cuál es la probabilidad de que una segunda muestra seleccionada tendría una media menor que 2.55?

- 13. Proporciones muestral y poblacional.** Suponga que en un suburbio de 12 345 personas, 6 523 de ellas se mudaron allí en los últimos cinco años. Usted encuesta a 500 personas y encuentra que 245 de las personas en su muestra se mudaron a los suburbios en los últimos cinco años.

- ¿Cuál es la proporción de la población de personas que se mudaron a los suburbios en los últimos cinco años?
- ¿Cuál es la proporción muestral de personas que se mudaron a los suburbios en los últimos cinco años?
- ¿Su muestra parece ser representativa de la población? Analice.

- 14. Proporciones muestral y poblacional.** Suponga que, en una escuela con 1 348 estudiantes, 137 estudiantes son zurdos. Usted encuesta a 100 estudiantes y encuentra que 11 de los estudiantes en su muestra son zurdos.

- ¿Cuál es la proporción de la población de estudiantes zurdos?
- ¿Cuál es la proporción muestral de estudiantes zurdos?
- ¿Su muestra parece ser representativa de la población? Analice.

- 15. Estimación de proporciones poblacionales.** Usted selecciona una muestra aleatoria de 150 personas en una convención médica a la que asistieron 1 608 personas. En su muestra encuentra que 73 personas han viajado del extranjero. Con base en este estadístico muestral, estima cuántas personas en la convención viajaron del extranjero. ¿Tendría más confianza de su estimación si hubiese muestreado a 300 personas? Explique.

- 16. Estimación de proporciones de población.** Una muestra aleatoria de personas se seleccionó de 74 512 personas que asistieron al juego del Súper Tazón en Miami. En la muestra, 420 personas están apoyando a los Osos de Chicago. Con base en este estadístico muestral, estime cuántas personas en el juego apoyan a los Osos de Chicago. ¿Estaría más confiado de su estimación si hubiese muestreado a 2 000 personas? Explique.

- 17. Distribución de proporciones muestrales.** Suponga que usted sabe que la distribución de las proporciones muestrales de no residentes en muestras de 200 estudiantes es normal con una media de 0.34 y una desviación estándar de 0.03. Suponga que selecciona una muestra aleatoria de 200 estudiantes y encuentra que la proporción de estudiantes no residentes en la muestra es 0.32.

- ¿A cuántas desviaciones estándar está la proporción muestral de la media de la distribución de las proporciones muestrales?
- ¿Cuál es la probabilidad de que una segunda muestra seleccionada tenga una proporción menor que 0.32?

- 18. Distribución de proporciones muestrales.** Suponga que usted sabe que la distribución de las proporciones muestrales de mujeres trabajadoras es normal con una media de 0.42 y una desviación estándar de 0.21. Suponga que usted selecciona una muestra aleatoria de empleados y encuentra que la proporción de mujeres en la muestra es 0.45.

- ¿A cuántas desviaciones estándar está la proporción muestral de la media de la distribución de proporciones muestrales?
- ¿Cuál es la probabilidad de que una segunda muestra seleccionada tenga una proporción mayor que 0.45?

- 19. Distribución muestral.** Un mariscal de campo lanzó una intercepción en su primer juego, 2 intercepciones en su segundo juego, 5 intercepciones en su tercer juego, y luego se retiró. Considere los valores de 1, 2 y 5 como la población. Suponga que se seleccionan muestras aleatorias de tamaño 2 con *reemplazo* de la población.

- Después de listar las 9 muestras diferentes, determine la media de cada muestra.
- ¿Cuál es la media de las medias muestrales del inciso a?
- ¿La media de la distribución muestral del inciso a es igual a la media de la población de los tres valores listados? ¿Esas medias *siempre* son iguales?

20. Distribuciones muestrales. A continuación está la población de los cinco presidentes de Estados Unidos que eran militares de profesión, junto con sus edades cuando tomaron posesión: Eisenhower, 62; Grant, 46; Harrison, 68; Taylor, 64 y Washington, 57. Suponga que muestras de tamaño 2 se seleccionan de manera aleatoria *con reemplazo* de la población de las cinco edades.

- Después de listar las 25 muestras posibles, determine la media de cada muestra.
- ¿Cuál es la media de las medias muestrales del inciso a?
- ¿La media de la distribución muestral (del inciso b) es igual a la media de la población de los cinco valores listados? ¿Esas medias *siempre* son iguales?

21. Formación de distribuciones muestrales. Cinco estados y sus áreas (en miles de millas cuadradas) se dan en la tabla siguiente. Considere estos cinco estados como la población completa de la cual se seleccionan muestras con reemplazo. Determine la distribución de las medias muestrales para muestras de tamaño $n = 1, 2, 3, 4$ y 5 . Determine la media de la distribución de medias muestrales en cada caso. Compare las medias de las distribuciones con la media de la población.

Estado	Área (miles de millas cuadradas)
Alabama	52
Connecticut	5
Georgia	60
Maine	33
Oregon	97

22. Formación de distribuciones muestrales. Al final de una sesión de práctica, los cinco titulares en un equipo de hockey habían anotado los siguientes números de puntos.

Jugador	Puntos
A	101
B	87
C	75
D	66
E	62

Determine la distribución de las medias muestrales para tamaños de muestra $n = 1, 2, 3, 4$ y 5 , donde las muestras se seleccionaron con reemplazo. En cada caso, determine la media de la distribución de medias muestrales. Compare

las medias de las distribuciones de medias muestrales con la media de la población.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 8 en www.aw.com/bbt.

23. Distribuciones de medias muestrales. Considere el conjunto grande de datos de horas que los estudiantes destinan al uso de búsqueda en internet, listado en esta sección en la página 337. Analice los métodos para la selección de una muestra aleatoria de esta población. Haga que cada persona en la clase seleccione una muestra aleatoria de $n = 10$ individuos de la población y determine la media de su muestra. Determine la media de las medias muestrales y compárela con la media de la población. Repita el proceso con muestras de tamaño $n = 20$. ¿Cómo se compara la media de las medias muestrales con la media de la población?

24. Distribuciones de proporciones muestrales. Considere el conjunto de datos de la página 341 que muestra 400 respuestas *sí* (Y) y *no* (N) a una pregunta de una encuesta (“¿Usted es propietario de un automóvil?”). Analice los métodos para la selección de una muestra aleatoria de esta población. Haga que cada persona en la clase seleccione una muestra aleatoria de $n = 10$ respuestas de la población y encuentre la proporción de respuestas *sí* (Y) para su muestra. Determine la media de las proporciones muestrales individuales y compárela con la proporción de la población de 0.6. Repita el proceso con muestras de tamaño $n = 20$. ¿Cómo la media de las proporciones muestrales se compara con la proporción de la población?



EN LAS NOTICIAS



25. Medias muestrales en las noticias. Encuentre una noticia o un reporte de investigación en el que la media muestral sea citada. Analice cómo es utilizada para estimar la media de la población.

26. Proporciones muestrales en las noticias. Encuentre una noticia o reporte de investigación en el que una proporción muestral sea citada. Analice cómo es usada para estimar una proporción de la población.

8.2 Estimación de medias poblacionales

En la sección 8.1 vimos cómo surgen las distribuciones de las medias muestrales y de las proporciones muestrales. También vimos que, si tenemos una sola muestra aleatoria de una población, su media muestral es nuestra mejor estimación de la media poblacional. Pero, ¿cuán buena es esta “mejor estimación”? En esta sección investigamos cómo hacer afirmaciones cuantitativas acerca de la incertidumbre cuando una media poblacional se infiere a partir de una media muestral. (En la sección 8.3 analizamos las proporciones poblacionales).

A propósito...

Las DDR de proteína en ocasiones se dan en gramos por kilogramo de peso corporal. Una cifra común para no atletas y mujeres que no están embarazadas es 0.8 gramos de proteínas por kilogramo de peso corporal.



Estimación de una media poblacional: los fundamentos

Las plantas y los animales de manera constante producen nuevas células, no sólo para crecer, sino también para reemplazar las células viejas. La producción de células nuevas requiere de proteínas, los bloques de construcción de los organismos vivos. Por esta razón, las proteínas son una parte esencial de nuestra dieta. ¿Cuántas proteínas debe comer por día? Muchas organizaciones nutricionales y agencias del gobierno proporcionan la dieta diaria recomendada (DDR) no sólo para proteína, sino también para vitaminas, minerales, carbohidratos y grasa. Estas DDR varían de acuerdo con la fuente, y cambian al paso del tiempo cuando se hace nueva investigación. La mayoría de las recomendaciones para consumo de proteínas son cercanas a alrededor de 55-60 gramos al día para hombres y 45-50 gramos al día para mujeres que no están embarazadas o en lactancia.

Dadas estas recomendaciones de cuántas proteínas *debemos* comer, es interesante preguntar en realidad cuántas proteínas *comen* los estadounidenses. La figura 8.8 muestra resultados parciales de una encuesta grande de los hábitos nutricionales de los estadounidenses (Tercera Encuesta Nacional de Salud y Examen Nutricional, o NHANES III, por sus siglas en inglés, realizada por el Centro Nacional para Estadísticas de Salud). El estudio original involucró aproximadamente 30 000 participantes que fueron encuestados en muchos aspectos diferentes de su salud y dieta. El histograma muestra la ingesta diaria promedio de proteínas, en gramos, para una muestra de $n = 267$ hombres tomados del estudio.

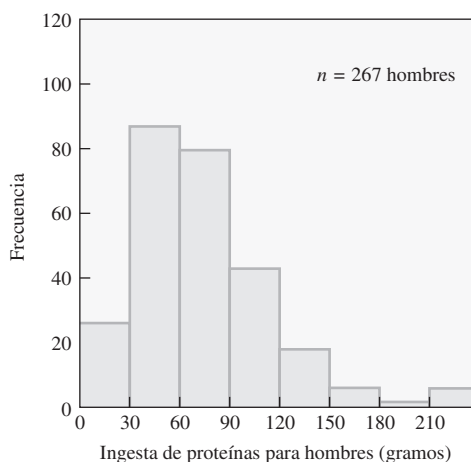


Figura 8.8 Histograma de ingesta diaria de proteínas para hombres de una muestra de $n = 267$ hombres. Fuente: Centro Nacional de Estadísticas de Salud.

Nuestro objetivo es utilizar la ingesta promedio de proteína para esta muestra de $n = 267$ hombres a fin de hacer una inferencia acerca de la ingesta promedio de proteína de todos los hombres estadounidenses. Los datos muestrales tienen una media de $\bar{x} = 77.0$ gramos y

una desviación estándar de $s = 58.6$ gramos. Como se analizó en la sección 8.1, cuando tenemos una sola muestra la media muestral es la mejor estimación de la media poblacional, μ . Sin embargo, no esperamos que la media muestral sea igual a la media poblacional, ya que existe probabilidad de que haya algún error de muestreo. Por tanto, para hacer una inferencia acerca de la media poblacional, necesitamos de alguna manera describir cuán bien esperamos que sea representada por la media muestral. El método más común para hacer esto es por medio de *intervalos de confianza*. Analizamos el uso de intervalos de confianza en el capítulo 1. Ahora estamos preparados para examinar cómo son calculados los intervalos de confianza.

La idea de un intervalo de confianza viene directamente del trabajo que hicimos con las distribuciones muestrales. Un cálculo preciso muestra que si la distribución de las medias muestrales es normal con media de μ , entonces 95% de todas las medias muestrales están dentro de 1.96 desviaciones estándar de la media poblacional; para nuestros propósitos en este libro, aproximaremos esto como 2 desviaciones estándar. Un **intervalo de confianza** es un rango de valores que probablemente contenga el valor verdadero de la media poblacional. En este libro sólo trabajamos con niveles del 95% de confianza, aunque en ocasiones se utilizan otros niveles de confianza (tal como 90% o 99%). Encontramos el intervalo de confianza del 95% a partir de la media muestral trabajando con el **margen de error**, como se define en el recuadro siguiente.

Intervalo de confianza del 95% para una media poblacional

El **margen de error** para el intervalo de confianza del 95% es

$$\text{margen de error} = E \approx \frac{2s}{\sqrt{n}}$$

donde s es la desviación estándar de la muestra. Encontramos el **intervalo de confianza del 95%** al sumar y restar el margen de error de la media muestral. Esto es, el intervalo de confianza del 95% va

$$\text{de } (\bar{x} - \text{margen de error}) \text{ a } (\bar{x} + \text{margen de error})$$

Podemos escribir este intervalo de confianza de manera más formal como

$$\bar{x} - E < \mu < \bar{x} + E$$

o de manera más breve como

$$\bar{x} \pm E$$

NOTA TÉCNICA

La fórmula precisa para el margen de error utiliza 1.96 en lugar de 2. Además, una consecuencia del teorema del límite central es que la distribución de las medias muestrales tiene una desviación estándar de $\sigma/\sqrt{n}$, pero s puede usarse en lugar de σ siempre que la muestra sea lo suficientemente grande (por lo común $n > 30$, pero el requerimiento del tamaño depende de la naturaleza de la distribución real de la población). La fórmula para el margen de error puede extenderse a otros niveles de confianza. Por ejemplo, para un nivel de confianza del 90%, reemplace el 2 en la fórmula por 1.645. Para un nivel de confianza del 99%, reemplace el 2 en la fórmula por 2.575.

La media muestral, $\bar{x}$, está en el centro del intervalo de confianza, el cual se extiende a una distancia igual al margen de error en cada dirección (figura 8.9). Decimos que estamos 95% confiados que el intervalo de confianza contiene al valor verdadero de la media poblacional. Este enunciado debe interpretarse de manera cuidadosa como sigue: si repitiésemos el proceso de obtener muestras y construir tales intervalos de confianza muchas veces, 95% de los intervalos de confianza contendría el valor de la media poblacional (μ) y 5% no lo incluiría.

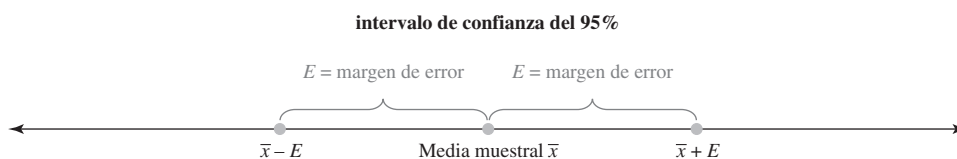


Figura 8.9 El intervalo de confianza del 95% se extiende una distancia igual al margen de error a los dos lados de la media muestral.

EJEMPLO 1 Cálculo del margen de error

Calcule el margen de error y encuentre el intervalo de confianza del 95% para muestra de ingesta muestral de $n = 267$ hombres, que tiene una media muestral de $\bar{x} = 77.0$ gramos y una desviación estándar de $s = 58.6$ gramos.

Solución El tamaño de la muestra es $n = 267$ y la desviación estándar para la muestra es $s = 58.6$, por lo que el margen de error es

$$E \approx \frac{2s}{\sqrt{n}} = \frac{2 \times 58.6}{\sqrt{267}} = 7.2$$

La media muestral es $\bar{x} = 77.0$ gramos, por lo que el intervalo de confianza del 95% se amplía aproximadamente desde $77.0 - 7.2 = 69.8$ gramos hasta $77.0 + 7.2 = 84.2$ gramos. Escribimos este resultado de manera más formal como

$$69.8 \text{ gramos} < \mu < 84.2 \text{ gramos}$$

o de manera más sencilla como 77.0 ± 7.2 gramos. Estamos confiados al 95% de que la media poblacional para la ingesta de proteínas de todos los estadounidenses hombres está entre 69.8 y 84.2 gramos. Es interesante notar que aunque el número menor en este intervalo de confianza (69.8 gramos) es mayor que la dieta diaria recomendada de proteína para hombres de 55-60 gramos, sugiere que el consumo real de proteínas es significativamente mayor que el recomendado.

NOTA TÉCNICA

Si utilizamos la fórmula precisa (con 1.96 en lugar de 2), encontramos que los límites del intervalo de confianza son 70.0 y 84.0.

Interpretación del intervalo de confianza

La figura 8.10 muestra una interpretación visual del intervalo de confianza. Imagine que tenemos, digamos, 20 muestras diferentes. Cada muestra tiene una media muestral diferente, $\bar{x}$, con un intervalo de confianza alrededor de la media muestral. Nunca podemos estar seguros de cuál intervalo contiene a la media de la población. Sin embargo, en promedio, 95% de las muestras, o 19 de las 20 muestras, tendrán un intervalo de confianza que capture la media poblacional verdadera. Un intervalo de confianza ocasional (alrededor de 5% de los generados) no captura la media poblacional verdadera.

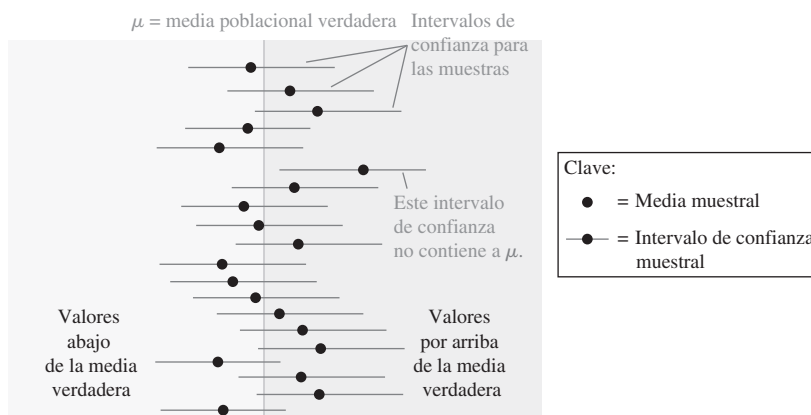


Figura 8.10 Esta figura ilustra la idea detrás de los intervalos de confianza. La línea vertical representa la media poblacional verdadera, μ . Cada una de las 20 líneas horizontales representa el intervalo de confianza de 95% para una muestra particular, con la media muestral marcada por el punto en el centro del intervalo de confianza. Con un intervalo de confianza de 95%, esperamos que 95% de todas las muestras darán un intervalo de confianza que contenga la media poblacional, como es el caso en esta figura, 19 de los 20 intervalos de confianza en realidad contienen la media poblacional. Esperamos que la media poblacional no estará dentro del intervalo de confianza en 5% de los casos; aquí, 1 de los 20 intervalos de confianza (el sexto de arriba hacia abajo) no contiene la media poblacional.

Sea cuidadoso —es fácil interpretar o establecer intervalos de confianza de manera incorrecta—. Es tentador afirmar que un intervalo de confianza del 95% significa que existe una probabilidad de 0.95 que la media poblacional caiga dentro del intervalo de confianza. Pero esto no es verdadero: aunque no podemos conocer la media poblacional real, μ , es un número fijo (no una variable), por lo que *cae* dentro del intervalo de confianza o *no cae* dentro del intervalo de confianza. No podemos hablar acerca de la probabilidad de que μ caiga dentro del intervalo de confianza. Para recapitular, la interpretación correcta de un intervalo de confianza del 95% es: *si repetimos el proceso de obtener muestras y construimos intervalos de confianza, a la larga 95% de los intervalos de confianza contendrán la media poblacional verdadera*.

EJEMPLO 2 Construcción de un intervalo de confianza

Un estudio determina que el tiempo promedio que destinan los alumnos de octavo grado viendo televisión es 6.7 horas a la semana, con un margen de error de 0.4 horas (para una confianza de 95%). Construya e interprete el intervalo de confianza del 95%.

Solución La mejor estimación de la media poblacional es la media muestral $\bar{x} = 6.7$ horas. Encontramos el intervalo de confianza sumando y restando el margen de error a la media muestral, por lo que el intervalo se extiende desde $6.7 - 0.4 = 6.3$ horas a $6.7 + 0.4 = 7.1$ horas. Por tanto, podemos asegurar con 95% de confianza que el tiempo promedio que observa televisión toda la población de alumnos de octavo grado está entre 6.3 y 7.1 horas, o

$$6.3 \text{ horas} < \mu < 7.1 \text{ horas.}$$

Si fueran tomadas 100 muestras aleatorias del mismo tamaño, esperaríamos que los intervalos de confianza de 95 de esas muestras contendrían la media poblacional.

EJEMPLO 3 Ingesta de proteína para mujeres

El estudio nutricional NHANES III también produjo datos sobre la ingesta de proteínas para mujeres. La figura 8.11 muestra un histograma para una muestra aleatoria de $n = 264$ mujeres (una parte pequeña del estudio completo). La media de estos datos es $\bar{x} = 59.6$ gramos y la desviación estándar es $s = 30.5$ gramos. Estime la media poblacional y proporcione un intervalo de confianza del 95%. Comente cómo se comparan estos valores con la dieta diaria recomendada (DDR) para mujeres de 45-50 gramos.

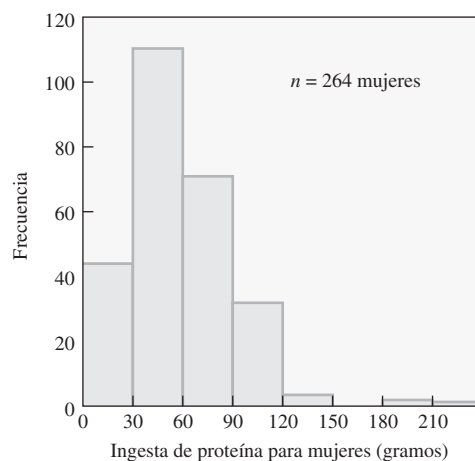


Figura 8.11 El histograma para la ingesta diaria de proteína para 264 mujeres.

Fuente: Centro Nacional de Estadísticas de Salud.

Solución La media muestral, $\bar{x} = 59.6$ gramos, es nuestra mejor estimación de la media poblacional. Para encontrar el intervalo de confianza al 95%, primero debemos encontrar el margen de error usando el tamaño de muestra ($n = 264$) y la desviación estándar muestral ($s = 30.5$ gramos):

$$E \approx \frac{2s}{\sqrt{n}} = \frac{2 \times 30.5 \text{ gramos}}{\sqrt{264}} = 3.8 \text{ gramos}$$

Por tanto, el intervalo de confianza del 95% va de $59.6 - 3.8 = 55.8$ gramos a $59.6 + 3.8 = 63.4$ gramos, o

$$55.8 \text{ gramos} < \mu < 63.4 \text{ gramos}$$

Podemos decir con 95% de confianza que la media poblacional —la ingesta media de proteína para mujeres— está dentro del intervalo de 55.8 gramos a 63.4 gramos. Concluimos que el consumo de proteína por las mujeres es mayor que la dieta diaria recomendada de 45-50 gramos por mujer.

UN MOMENTO DE REFLEXIÓN

Recuerde que la desviación estándar de los datos de la ingesta de proteína para la muestra de hombres fue $s = 58.6$ gramos —casi el doble de la desviación estándar para la muestra de mujeres ($s = 30.5$ gramos)—. ¿Esta diferencia cómo afecta los márgenes de error? Haga conjeturas de por qué las desviaciones estándar son diferentes.

EJEMPLO 4 Producción de basura

Un estudio realizado por el Proyecto de Basura en la Universidad de Arizona analizó el contenido de basura desechada por $n = 62$ hogares; los hogares variaron en tamaño de 2 a 11 miembros. El histograma en la figura 8.12 muestra la producción de basura semanal (en libras) para los hogares en la muestra. (El estudio completo proporciona el desglose de basura para varias categorías). La media para la muestra es $\bar{x} = 27.4$ libras y la desviación estándar es $s = 12.5$ libras. Estime la media de la población para la producción de basura semanal con un intervalo de confianza del 95%. Comente sobre las conclusiones del estudio.

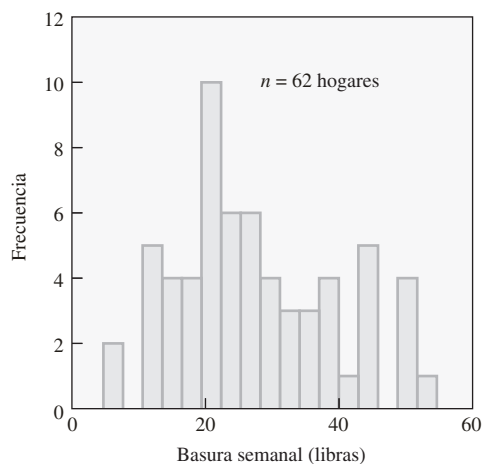


Figura 8.12 Histograma para la producción total de basura para $n = 62$ hogares.

Solución La media muestral, $\bar{x} = 27.4$ libras, es nuestra mejor estimación de la media poblacional. Utilizamos la fórmula del margen de error para encontrar que

$$E \approx \frac{2s}{\sqrt{n}} = \frac{2 \times 12.5 \text{ libras}}{\sqrt{62}} = 3.2 \text{ libras}$$

A propósito...

De acuerdo con la información del Proyecto de Basura, el componente principal de la mayoría de la basura de los hogares es papel (constituyendo de un tercio a un medio de la basura total), seguido de alimento y vidrio. Además, existe una correlación significativa entre el peso del plástico desechado y el tamaño de un hogar, por lo que el plástico desechado podría usarse para estimar la población de una región.

Por tanto, el intervalo de confianza del 95% va de $27.4 - 3.2 = 24.2$ libras a $27.4 + 3.2 = 30.6$ libras, o

$$24.2 \text{ libras} < \mu < 30.6 \text{ libras}$$

Con una confianza del 95%, podemos asegurar que el intervalo de 24.2 a 30.6 libras contiene la cantidad media de basura desechada por hogares estadounidenses en una semana. Se podrían hacer dos observaciones acerca de esta conclusión. Primera, el amplio rango de tamaños de hogares produce variación significativa en los datos. Esta variación está reflejada en una gran desviación estándar que produce un margen de error grande. La conclusión podría ser más significativa si estuviese dada en términos de hogares del mismo tamaño. Segunda, el tamaño relativamente pequeño (62 hogares) de la muestra también contribuye al gran margen de error. Una conclusión más confiable podría obtenerse de una muestra más grande.

EJEMPLO 5 Temperatura corporal

Un estudio realizado por investigadores de la Universidad de Maryland registró las temperaturas del cuerpo para $n = 106$ sujetos. La media muestral del conjunto de datos es $\bar{x} = 98.20^\circ\text{F}$ y la desviación estándar para la muestra es $s = 0.62^\circ\text{F}$. Estime la media de la temperatura corporal para la población con un intervalo de confianza del 95%.

Solución La media muestral, $\bar{x} = 98.20^\circ\text{F}$, es nuestra mejor estimación de la media de la temperatura corporal de la población. El margen de error es

$$E \approx \frac{2s}{\sqrt{n}} = \frac{2 \times 0.62^\circ\text{F}}{\sqrt{106}} = 0.12^\circ\text{F}$$

Por tanto, el intervalo de confianza del 95% varía desde $98.20^\circ\text{F} - 0.12^\circ\text{F} = 98.08^\circ\text{F}$ hasta $98.20^\circ\text{F} + 0.12^\circ\text{F} = 98.32^\circ\text{F}$, o

$$98.08^\circ\text{F} < \mu < 98.32^\circ\text{F}$$

Interpretamos este resultado como sigue: si seleccionásemos muchas muestras de tamaño $n = 106$ y calculásemos los intervalos de confianza de todas las muestras, esperamos que 95% de los intervalos de confianza contendrían la media poblacional verdadera. Observe que la temperatura corporal media citada para los humanos (98.6°F) *no* está contenida en el intervalo de confianza. Con base en esta muestra, es probable que la temperatura corporal media aceptada esté equivocada.

Selección del tamaño de la muestra

Al planear encuestas estadísticas y experimentos, con frecuencia conocemos por anticipado el margen de error que nos gustaría alcanzar. Por ejemplo, podríamos querer estimar el costo de un automóvil nuevo dentro de \$200. Podemos estimar el tamaño de la muestra, n , que se necesita para asegurar este margen de error ($E \approx 2s/\sqrt{n}$) para n . Con un poco de álgebra, encontramos

$$n \approx \left(\frac{2s}{E} \right)^2$$

La fórmula exacta para el tamaño de muestra utiliza la desviación estándar poblacional, σ , en lugar de s . En la práctica, es muy raro que conozcamos la desviación estándar de la población, ya que estudiamos sólo muestras. Por tanto, para usar la fórmula exacta del tamaño de muestra, por lo regular estimamos la desviación estándar poblacional con base en estudios previos, estudios piloto o suposiciones bien fundamentadas. El tamaño de cualquier muestra real debe ser un número entero, por lo que redondeamos el resultado de la fórmula para el tamaño de la muestra *hacia arriba* para el entero más cercano. Cualquier muestra mayor que este tamaño nos da un margen de error tan pequeño o menor al que buscamos.

NOTA TÉCNICA

La fórmula dada para el tamaño de la muestra supone que estamos trabajando con un nivel de confianza del 95%; una fórmula más precisa sería utilizar 1.96 en lugar de 2, como se analizó anteriormente. Para los métodos de esta sección, si la desviación estándar poblacional no es conocida, podemos utilizar la desviación estándar muestral, s , como una estimación de ese valor. Sin embargo, los resultados serán buenos sólo si el tamaño de la muestra es grande. Si se toma una muestra pequeña de una población no normal, el uso de s en lugar de σ podría llevar a resultados muy malos.

Selección del tamaño de muestra correcto

Para estimar la media poblacional con un margen de error especificado de a lo más E , el tamaño de la muestra debe ser al menos

$$n = \left(\frac{2\sigma}{E} \right)^2$$

donde σ es la desviación estándar de la población (con frecuencia estimada mediante la desviación estándar muestral, s).

EJEMPLO 6 Costo medio de casas

Usted quiere estudiar los costos de casas en el país con un muestreo de las ventas recientes de casas en varias regiones (representativas). Su objetivo es proporcionar una estimación del intervalo de confianza del 95% del costo de las casas. Estudios previos sugieren que la desviación estándar poblacional es de alrededor de \$7200. ¿Qué tamaño (mínimo) de muestra debe usarse para asegurar que la media muestral está dentro de

- a. \$500 de la media poblacional verdadera?
- b. \$100 de la media poblacional verdadera?

Solución

- a. Con $E = \$500$ y σ estimada como \$7200, el tamaño mínimo de muestra que cumple los requisitos es

$$n = \left(\frac{2\sigma}{E} \right)^2 = \left(\frac{2 \times 7200}{500} \right)^2 = 28.8^2 = 829.4$$

Puesto que el tamaño de muestra debe ser entero, concluimos que la muestra debe incluir *por lo menos* 830 precios.

- b. Con $E = \$100$ y $\sigma = \$7200$, el tamaño mínimo de muestra que cumple con los requisitos es

$$n = \left(\frac{2\sigma}{E} \right)^2 = \left(\frac{2 \times 7200}{100} \right)^2 = 144^2 = 20736$$

Observe que para disminuir el margen de error por un factor de 5 (de \$500 a \$100), debemos aumentar el tamaño de la muestra en un factor de 25. Por lo cual, obtener una mayor precisión, por lo general, viene acompañado de un alto costo.

UN MOMENTO DE REFLEXIÓN

Si usted decide que quiere un margen de error más pequeño para un intervalo de confianza, ¿debe aumentar o disminuir el tamaño de la muestra? Explique.

Sección 8.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Intervalo de confianza.** Con base en una muestra aleatoria de medidas de mujeres, construimos este intervalo de confianza del 95% que estima la media de nivel de colesterol de todas las mujeres:

$$183.3 \text{ miligramos} < \mu < 298.5 \text{ miligramos.}$$

Interprete este intervalo de confianza.

- Margen de error.** Con base en una muestra aleatoria de mediciones de presión sanguínea de hombres, la media muestral es 73.2 mm Hg y el margen de error para un intervalo de confianza del 95% es 2.8 mm Hg. Identifique el intervalo de confianza.
- Intervalos de confianza en la media.** A continuación está un enunciado común hecho para la media. “Con base en un estudio reciente, las monedas de un centavo pesan un promedio de 2.5 gramos con un margen de error de 0.006 gramos”. ¿Qué parte importante y relevante de información está omitida en el enunciado? ¿Es correcto utilizar la palabra “promedio”?
- Tamaño de la muestra.** El Examen Nacional de Salud incluye la medición de alrededor de 25 000 personas y los resultados son usados para estimar valores de varias medias poblacionales. ¿Es válido criticar esta encuesta porque el tamaño de la muestra es sólo de alrededor de 0.01% de la población de todos los estadounidenses? Explique.

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Estimación de la media.** Se obtuvo una encuesta de 200 lectores, elegidos de manera aleatoria, y se utilizó para determinar el intervalo de confianza de \$47 200.
- Margen de error.** El ingreso medio de maestros de matemáticas universitarios se estimó en \$48 213 con un margen de error de 5%.
- Margen de error.** Cuando datos muestrales son usados para estimar el valor de una media poblacional, el margen de error aumenta conforme el tamaño de la muestra aumenta.
- Mejor estimación.** Cuando datos muestrales fueron usados para estimar el valor del peso medio de todos los centavos, se obtuvo este intervalo de confianza del 95%:

$$2.495 \text{ gramos} < \mu < 2.505 \text{ gramos}$$

Con base en ese resultado, la mejor estimación con un solo valor es 2.500 gramos.

Conceptos y aplicaciones

Determinación de márgenes de error e intervalos de confianza. Para los ejercicios 9 al 12 suponga que las medias poblacionales tienen que estimarse a partir de las muestras descritas. En cada caso utilice los resultados muestrales para aproximar el margen de error y un intervalo de confianza del 95%.

- Tamaño muestral = 100, media muestral = 75.0, desviación estándar muestral = 10.0
- Tamaño muestral = 64, media muestral = 2.50, desviación estándar muestral = 1.08
- Tamaño muestral = 1 068, media muestral = \$46 205, desviación estándar muestral = \$24 000
- Tamaño muestral = 692, media muestral = 155 kilogramos, desviación estándar muestral = 27 kilogramos

Determinación de márgenes de error e intervalos de confianza. Para los ejercicios 13 al 16 suponga que queremos construir un intervalo de confianza del 95% de una media poblacional. Determine una estimación del tamaño de la muestra necesaria para obtener el margen de error especificado para el intervalo de confianza del 95%. La desviación estándar muestral está dada.

- Margen de error = 1.0 centímetros, desviación estándar = 3.5 centímetros
- Margen de error = 2.0 centigramos, desviación estándar = 16 centigramos
- Margen de error = 0.01 kilómetros, desviación estándar = 0.11 kilómetros
- Margen de error = 24 milímetros, desviación estándar = 416 milímetros
- Tamaño de muestra para una encuesta de televisión.** Nielsen Media Research desea estimar el número medio de horas que estudiantes de preparatoria destinan a ver televisión en un día entre semana. Se desea un margen de error de 0.25 horas. Estudios previos sugieren que una desviación estándar poblacional razonable es 1.7 horas. Estime el tamaño mínimo de muestra necesario para obtener la media poblacional con la precisión establecida.
- Tamaño de muestra para casas.** Una encuesta del gobierno realizada para estimar el precio medio de casas en una gran área metropolitana se diseñó para que tuviese un margen de error de \$10 000. Estudios piloto sugieren que la desviación estándar poblacional es \$65 500. Estime el tamaño de muestra mínimo que se necesita para calcular la media poblacional con la precisión establecida.
- Tamaño de muestra para el CI medio de estudiantes de estadística.** La prueba de CI Wechsler está diseñada para que la media sea 100 y la desviación estándar sea 15 para la población de adultos normales. Determine el tamaño de

muestra necesario para estimar la calificación media del CI de residentes en Delaware. Queremos 95% de confianza de que nuestra media esté dentro de 2 puntos del CI de la media verdadera. Luego suponga que $\sigma = 15$ y determine el tamaño de muestra necesario.

- 20. Tamaño de muestra para la estimación del ingreso.** Una economista necesita estimar el ingreso medio anual del primer año de trabajo para graduados universitarios que han tenido el gran acierto de tomar un curso de estadística. ¿Cuántos ingresos debe encontrar, si ella quiere tener confianza de 95% que la media muestral está dentro de \$500 de la media población verdadera? Suponga que un estudio previo ha revelado que para tales ingresos, $\sigma = \$6\,250$.

- 21. Peso de monedas de 25 centavos.** Usted quiere estimar el peso medio de las monedas de 25 centavos en circulación. Una muestra de 40 de ellas tiene un peso medio de 5.639 gramos y una desviación estándar de 0.062 gramos. Utilice un solo valor para estimar el peso medio de todas las monedas de 25 centavos. Además, encuentre el intervalo de confianza del 95%.

- 22. Calificaciones del SAT.** Una muestra de 250 estudiantes de primer año en una gran universidad tiene una calificación media del SAT de 1722 con una desviación estándar de 280. Utilice un solo valor para estimar la calificación media del SAT para toda la clase de primer año. Además, determine el intervalo de confianza del 95%.

- 23. Tiempo para la terminación de estudios.** Datos del Centro Nacional para Estadísticas de Educación de 4 400 graduados universitarios muestran que el tiempo medio para graduarse de licenciatura es 5.15 años con una desviación estándar de 1.68 años. Utilice un solo valor para estimar el tiempo medio para graduarse de todos los egresados universitarios. Además, determine el intervalo de confianza del 95%.

- 24. Producción de basura.** Con base en una muestra de 62 hogares, el peso medio de plástico desechado es 1.91 libras y la desviación estándar es 1.07 libras (datos del Proyecto de Basura de la Universidad de Arizona). Utilice un solo valor para estimar el peso medio de plástico desechado por todos los hogares. Además, determine el intervalo de confianza del 95%.

- 25. Peso de osos.** La salud de la población de osos en el Parque Nacional de Yellowstone es monitoreada mediante mediciones periódicas tomadas a osos anestesiados. Una muestra de los pesos de tales osos se da a continuación. Determine un intervalo de confianza del 95% que estime la media de la población de los pesos de tales osos.

80	344	416	348	166	220	262	360	204
144	332	34	140	180	105	166	204	26
120	436	125	132	90	40	220	46	154
116	182	150	65	356	316	94	86	150

- 26. Nivel de cotinina en fumadores.** Cuando la gente fuma, la nicotina que ellos absorben se convierte en cotinina, que puede medirse. Una muestra de niveles de cotinina de 40

fumadores se lista a continuación. Determine un intervalo de confianza del 95% para estimar la media del nivel de cotinina de todos los fumadores.

1	0	131	173	265	210	44	277
32	3	35	112	477	289	227	103
222	149	313	491	130	234	164	198
17	253	87	121	266	290	123	167
250	245	48	86	284	1	208	173

- 27. Tamaño de una familia.** Usted selecciona una muestra aleatoria de $n = 31$ familias en su vecindario y determina los tamaños de familia siguientes (número de personas en la familia):

2	3	6	5	4	2	3	3	1	2	3
2	3	4	5	3	1	3	3	4	7	3
2	3	2	2	3	4	1	5	2		

- ¿Cuál es el tamaño medio de las familias para la muestra?
- ¿Cuál es la desviación estándar para la muestra?
- ¿Cuál es la mejor estimación para el tamaño medio de las familias para la población de todas las familias estadounidenses?
- ¿Cuál es el intervalo de confianza del 95% para la estimación?
- Comente sobre la fiabilidad de la estimación.

- 28. Aparatos televisores.** A una muestra de $n = 31$ hogares se les preguntó sobre el número de televisores en la casa. Las respuestas fueron las siguientes.

1	0	2	3	2	3	4	2	1	1	2
4	3	2	3	3	0	1	0	1	3	2
4	3	2	1	4	0	1	2	3		

- ¿Cuál es el número medio de televisores para la muestra?
- ¿Cuál es la desviación estándar para la muestra?
- ¿Cuál es la mejor estimación para el número medio de televisores para la población de todos los hogares estadounidenses?
- ¿Cuál es el intervalo de confianza del 95% para la estimación?
- Comente sobre la fiabilidad de la estimación.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 8 en www.aw.com/bbt.

- 29. Antigüedad de automóviles.** Suponga que usted quiere estimar la antigüedad media de automóviles conducidos por estudiantes en su escuela. Un estudio previo muestra que la desviación estándar de esas edades es aproximadamente 3.7 años. ¿Cuántas edades de automóviles deben seleccionarse de manera aleatoria para tener una confianza de 95% que su media muestral está dentro de 1 año de la media poblacional? Usando ese tamaño de muestra recolecte sus propios datos muestrales, que consisten en las edades

de automóviles manejados por estudiantes de su escuela. Luego utilice los métodos de esta sección para construir un intervalo de confianza del 95%. Escriba un enunciado que resuma sus resultados.

30. Compañías de encuestas. Tres compañías líderes en encuestas públicas son Gallup Organization, Harris Poll y Yankelovich Partners. Visite sus sitios web. Describa la historia de cada organización y los servicios de encuestas que proporcionan. ¿Cuál compañía tiene la mejor descripción de sus métodos de encuestas?

31. Encuestas de cadenas de televisión. Todas las cadenas importantes de televisión realizan encuestas regularmente sobre varios temas. Visite el sitio web de por lo menos una de las cadenas principales y reúna los resultados de una encuesta particular que involucre la estimación de una media poblacional. Asegúrese de incluir toda la información que se da acerca del tamaño de la muestra, margen de error e intervalos de confianza.

32. Corrupción internacional. Transparencia Internacional utiliza encuestas para determinar el Índice Pagadores de Soborno y el Índice de Percepción de Corrupción que miden el grado de corrupción en muchos países del mundo. Visite el sitio web de Transparencia Internacional y revise la extensa documentación que describe los métodos usados por esta organización. Analice los resultados y su validez.

EN LAS NOTICIAS

33. Estimación de medias poblacionales. Encuentre una noticia en la que una media poblacional se estima a partir de una muestra. El artículo debe incluir un margen de error y/o un intervalo de confianza. Analice los métodos usados en el estudio y cómo se llegó a las conclusiones.

8.3 Estimación de proporciones poblacionales

En esta sección volvemos nuestra atención para estimar *proporciones* poblacionales. Muchos sondeos de opinión y encuestas conocidas dependen de las técnicas que analizaremos. Por ejemplo, los índices de audiencia Nielsen estiman la proporción de la población que sintoniza cierto programa de radio y televisión, las cifras de desempleo mensual editadas por la Oficina de Estadísticas del Trabajo son estimadas con base en la proporción de estadounidenses que están desempleados, y las encuestas de opinión que dominan la política estadounidense estiman la proporción de la población que apoya a un candidato o medida particular.

Los fundamentos de la estimación de una proporción poblacional

La Oficina de Estadísticas de Trabajo estima la tasa de desempleo a partir de una encuesta mensual a 60 000 hogares (vea el ejemplo 2 en la sección 1.1). La tasa de desempleo para esta muestra es la *proporción muestral* (la proporción de personas en la muestra que está desempleada), denotada $\hat{p}$. La proporción muestral es la mejor estimación para la *proporción poblacional* (la proporción de gente en la población que está desempleada), denotada p .

Al igual que con las medias poblacionales, una estimación de la proporción poblacional puede entenderse mejor si decimos algo acerca de su precisión. Otra vez, usamos márgenes de error e intervalos de confianza. El único cambio de la estimación de las medias poblacionales estriba en la definición del margen de error, que está dado en el recuadro siguiente. La figura 8.13 muestra cómo interpretamos el intervalo de confianza.

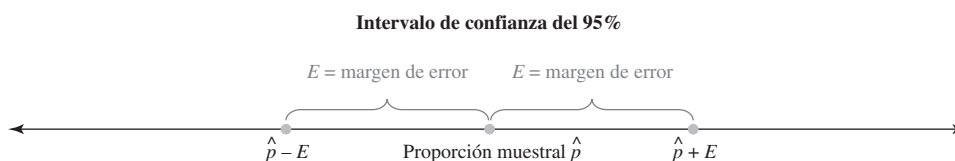


Figura 8.13 El intervalo de confianza se extiende una distancia igual al margen del error a cada lado de la proporción muestral, $\hat{p}$.

NOTA TÉCNICA

La fórmula precisa para el margen de error utiliza 1.96 en lugar de 2. La fórmula puede extenderse a otros niveles de confianza. Para un nivel de confianza del 90%, reemplace el 2 en la fórmula por 1.645. Para un nivel de confianza de 99%, reemplace el 2 en la fórmula por 2.575. La fórmula del margen de error dada aquí necesita que $np \geq 5$ y $n(1-p) \geq 5$, condiciones que por lo regular, con facilidad, se cumplen en la práctica. (Técnicas tratadas en textos más avanzados funcionan para muestras de tamaño pequeño).

Apropósito...

A continuación está cómo la Oficina de Estadísticas del Trabajo describe incertidumbre y su encuesta de desempleo. "Una muestra no es un conteo total y la encuesta no puede producir los mismos resultados que los que se hubiesen obtenido de la entrevista a toda la población. Pero las posibilidades son 90 en 100 de que la estimación de desempleo de la muestra esté dentro de 230 000 de la cifra obtenible de un censo total. Ya que los totales de desempleo mensual tienen un rango de alrededor de 6 a 11 millones en años recientes, el error posible como resultado del muestreo no es tan grande para distorsionar la imagen total de desempleo".

Intervalo de confianza del 95% para una proporción poblacional

Para una proporción poblacional el **margen de error** para el intervalo de confianza del 95% es

$$E \approx 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

donde $\hat{p}$ es la proporción muestral. El **intervalo de confianza del 95%** va

desde $\hat{p} - \text{margen de error}$ hasta $\hat{p} + \text{margen de error}$

Podemos escribir este intervalo de confianza de manera más formal como

$$\hat{p} - E < p < \hat{p} + E$$

EJEMPLO 1 Tasa de desempleo

La Oficina de Estadísticas del Trabajo encuentra 2 160 personas desempleadas en una muestra de $n = 60\,000$ personas. Estime la tasa de desempleo poblacional y proporcione un intervalo de confianza del 95%.

Solución La proporción muestral es la tasa de desempleo para la muestra:

$$\hat{p} = \frac{2\,160}{60\,000} = 0.036$$

Éste es el mejor estimador para la tasa de desempleo poblacional. El margen de error es

$$E \approx 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2\sqrt{\frac{0.036(1-0.036)}{60\,000}} = 0.0015$$

(La aproximación es válida puesto que el tamaño de la muestra es grande). El intervalo de confianza del 95% va desde $0.0360 - 0.0015 = 0.0345$ a $0.0360 + 0.0015 = 0.0375$, o

$$0.0345 < p < 0.0375$$

Podemos tener un 95% de confianza de que el intervalo de 3.45% a 3.75% contiene la tasa de desempleo verdadera para la población. Interpretamos este resultado como sigue: si calculamos los intervalos de confianza para muchas muestras de tamaño $n = 60\,000$, debemos esperar que 95% de los intervalos de confianza contengan a la proporción poblacional verdadera.

EJEMPLO 2 Índices de audiencia de televisión Nielsen

Los índices de audiencia Nielsen para televisión utilizan una muestra aleatoria de hogares. Una encuesta Nielsen dio como resultado para un juego de la Copa Mundial de Fútbol Femenil una estimación de 72.3% de la audiencia total. Suponiendo que la muestra consiste en $n = 5\,000$ hogares seleccionados de manera aleatoria, determine el margen de error y el intervalo de confianza del 95% para esta estimación.

Solución La proporción muestral, $\hat{p} = 72.3\% = 0.723$, es la mejor estimación de la proporción poblacional. El margen de error es

$$E \approx 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2\sqrt{\frac{0.723(1-0.723)}{5\,000}} = 0.013$$

El intervalo de confianza del 95% es $0.723 - 0.013 < p < 0.723 + 0.013$, o

$$0.710 < p < 0.736$$

Con una confianza del 95%, concluimos que entre 71.0% y 73.6% de la audiencia total observó el juego de fútbol de la Copa Mundial femenil.

EJEMPLO 3 Encuestas de Gallup

Desde 1935 la Organización Gallup ha sido líder en la medición y análisis de actitudes, opiniones y comportamiento de personas. Aunque Gallup es mejor conocida por la Encuesta Gallup, la compañía también proporciona investigación de mercado y de administración para grandes compañías. En una encuesta reciente a 1016 adultos seleccionados aleatoriamente se les preguntó:

Con lo que sabe, el antiguo jugador de las ligas mayores, Pete Rose, es no elegible para el Salón de la Fama del béisbol dados los cargos que tuvo de apuestas en juegos de béisbol. ¿Usted considera que debe o no ser elegible para admitirlo al Salón de la Fama?

Entre los encuestados, 59% consideró que Pete Rose debe ser elegible. En otra muestra más pequeña de 628 adultos identificados ellos mismos como “fanáticos del béisbol”, 62% consideró que él debía ser elegible para su admisión al Salón de la Fama. Encuentre el margen de error y el intervalo de confianza para cada muestra. La encuesta de los fanáticos de béisbol citaba un margen de error de “no más de 5 puntos porcentuales”. ¿Esta afirmación es consistente con el tamaño de la muestra?

Solución Una muestra con $n = 1016$ encuestados y una proporción muestral de $\hat{p} = 0.59$ tiene un margen de error de

$$E \approx 2\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 2\sqrt{\frac{0.59(1 - 0.59)}{1016}} = 0.031$$

o alrededor de 3 puntos porcentuales. Por tanto, el intervalo de confianza del 95% va desde $59 - 3 = 56\%$ hasta $59 + 3 = 62\%$.

Una muestra con $n = 628$ encuestados y una proporción muestral de $\hat{p} = 0.62$ tiene un margen de error de

$$E \approx 2\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 2\sqrt{\frac{0.62(1 - 0.62)}{628}} = 0.039$$

o alrededor de 4 puntos porcentuales. El intervalo de confianza del 95% va desde $62 - 4 = 58\%$ a $62 + 4 = 66\%$. Observe que el margen de error citado de “no más de 5 puntos porcentuales” es, de hecho, una sobreestimación.

Selección del tamaño de la muestra

Los diseñadores de encuestas y sondeos de opinión con frecuencia especifican cierto nivel de precisión de sus resultados. Por ejemplo, podía ser deseable estimar una proporción poblacional con un intervalo de confianza del 95% y un margen de error de no más de 1.5 puntos porcentuales. En tales situaciones es necesario determinar cuán grande debe ser la muestra para garantizar esta precisión. Mientras utilicemos un nivel de confianza del 95%, podemos trabajar con esta fórmula simplificada de aproximación para el margen de error:

$$E \approx \frac{1}{\sqrt{n}}$$

Esta fórmula proporciona una estimación conservadora (mayor que la necesaria) para el margen de error. Al despejar n se obtiene el tamaño de la muestra necesario para obtener un margen de error E :

$$n \approx \frac{1}{E^2}$$

Cualquier tamaño de muestra igual o mayor que este valor será suficiente.

NOTA TÉCNICA

Puede deducir la fórmula $E \approx 1/\sqrt{n}$ de la fórmula más precisa dada anteriormente para el margen de error, reemplazando el producto $\hat{p}(1 - \hat{p})$ por su valor máximo posible de 0.25. Esta aproximación sobreestima el margen de error real y es más precisa cuando p es cercana a 0.5.

Selección del tamaño correcto de muestra

Para estimar una proporción poblacional con un grado de 95% de confianza y un margen de error especificado de E , el tamaño de la muestra debe ser por lo menos

$$n = \frac{1}{E^2}$$

EJEMPLO 4 Mínimo tamaño de la muestra para una encuesta

Usted planea estimar la proporción de estudiantes en su campus que regularmente llevan un teléfono celular. ¿Cuántos estudiantes deben estar en la muestra si quiere un margen de error (con una confianza del 95%) de no más de 4 puntos porcentuales?

Solución Observe que 4 puntos porcentuales significan un margen de error de 0.04. Con base en la fórmula, el tamaño mínimo de la muestra es

$$n = \frac{1}{E^2} = \frac{1}{0.04^2} = 625$$

Debe encuestar por lo menos 625 estudiantes.

EJEMPLO 5 Encuesta Yankelovich

Yankelovich Partners es una compañía de investigación de opinión pública y de investigación de mercado. La compañía hace encuestas regularmente para la revista *Time* y las noticias de CNN. Los resultados de sus encuestas pueden encontrarse en la publicación mensual *Yankelovich Monitor*. Una encuesta reciente concluyó que 61% de todos los hogares tienen una computadora, con un margen de error de 3.5 puntos porcentuales. Aproximadamente, ¿qué tamaño de muestra debe haberse utilizado en esta encuesta?

Solución Un margen de error de 3.5 puntos porcentuales (0.035) podría ser obtenido con un tamaño de muestra de

$$n = \frac{1}{E^2} = \frac{1}{0.035^2} = 816.3$$

Puesto que debemos redondear al siguiente entero mayor, concluimos que fueron encuestados aproximadamente 817 hogares.

Sección 8.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Intervalo de confianza.** Con base en una encuesta realizada por ICR Survey Research Group, se obtuvo el siguiente intervalo de confianza del 95% para la proporción poblacional p : $0.457 < p < 0.551$. Interprete ese intervalo de confianza.
- Margen de error.** En una muestra de 8411 accidentes en el aterrizaje por aeronaves de aviación general, la proporción de accidentes en los que los pilotos murieron es 0.052. Cuando un intervalo de confianza del 95% es construido para la proporción poblacional de todos los accidentes, el margen de error se encontró que fue 0.005. Identifique el intervalo de confianza.

- Intervalos de confianza en los medios.** A continuación está un enunciado común hecho en los medios: “Con base en la encuesta, de todos los probables votantes 38% votarán por Díaz. El margen de error para esta encuesta es 3 puntos porcentuales”. ¿Qué parte importante y relevante de información es omitida en ese enunciado?
- Muestreo de conveniencia.** La presidenta de la Sociedad de Alumnos en la Universidad de Newport dirige una encuesta a 3250 estudiantes. Con base en los resultados, ella construye el intervalo de confianza del 95% para la proporción de todos los estudiantes universitarios de Estados Unidos, que se han emborrachado. Ella asegura que el tamaño grande de la muestra compensa el hecho que todos sus sujetos sean de la misma universidad. ¿Esta afirmación es válida? ¿Por qué sí o por qué no?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Intervalo de confianza.** El *Kingston Chronicle* publica un artículo que establece que los datos de la encuesta fueron usados para desarrollar un intervalo de confianza de 0.45.
6. **Intervalo de confianza en los medios.** El *Kingston Chronicle* publica un artículo en el que establece que, con base en los resultados de la encuesta, 82% de los residentes del condado de Orange se oponen al aumento en los impuestos, con un margen de error de 4 puntos porcentuales. Un lector dice que esto puede expresarse como el intervalo de confianza $0.78 < p < 0.86$.
7. **Tamaño de la muestra.** El intervalo de confianza del 95% de $0.200 < p < 0.400$ tiene como base un tamaño de muestra de 500. Si el tamaño de muestra se aumenta, el intervalo de confianza se hará más pequeño (más angosto).
8. **Tamaño de la muestra.** Un reportero del *Kingston Chronicle* asegura que cualquier buen intervalo de confianza debe tener como base una muestra que sea al menos del 5% del tamaño de la población.

Conceptos y aplicaciones

Márgenes de error e intervalos de confianza. Para los ejercicios del 9 al 12 suponga que las proporciones poblacionales son estimadas a partir de las muestras descritas. En cada caso determine el margen de error aproximado y el intervalo de confianza del 95%.

9. Tamaño de la muestra = 100, proporción muestral = 0.25
10. Tamaño de la muestra = 400, proporción muestral = 0.40
11. Tamaño de la muestra = 1068, proporción muestral = 0.228
12. Tamaño de la muestra = 1492, proporción muestral = 0.377

Tamaño de la muestra. Para los ejercicios del 13 al 16 estime el tamaño mínimo de la muestra necesario para obtener el margen de error dado.

13. $E = 0.01$
14. $E = 0.03$
15. $E = 0.06$
16. $E = 0.025$
17. **Índices de audiencia Nielsen.** Nielsen Media Research utiliza muestras de 5000 hogares para clasificar los programas de televisión. Nielsen reportó que *60 Minutes* tenía 15% de la audiencia de televisión. ¿Cuál es un intervalo de confianza del 95% para este resultado?
18. **Índices de audiencia Nielsen.** Repita el ejercicio 17 suponiendo que el tamaño de la muestra se duplica a 10000. Dado que el mayor costo y esfuerzo de realizar la encuesta

Nielsen sería duplicado, ¿este incremento en el tamaño de la muestra parece estar justificado por el aumento de confiabilidad?

19. **Riesgo de atletas.** Un estudio realizado por investigadores en la Universidad de Alfred concluyó que 80% de todos los atletas estudiantes en Estados Unidos han corrido alguna forma de riesgo. El estudio está basado en respuestas de 1400 atletas. ¿Cuál es el margen de error y el intervalo de confianza del 95% para el estudio?
20. **Opiniones de estudiantes.** Una encuesta anual de estudiantes universitarios de primer año, realizada por el Instituto de Investigación de Educación Superior en la UCLA, pregunta a aproximadamente 276 000 estudiantes acerca de sus actitudes en una variedad de temas. De acuerdo con una encuesta reciente, 51% de los estudiantes de primer año creen que el aborto debe ser legal (bajó de 65% en 1990) y 40% creen que el sexo casual es aceptable (bajó de 50% en 1975). ¿Cuáles son los márgenes de error y los intervalos de confianza del 95% para estas estimaciones?
21. **Hábitos en estados.** Los centros para Control y Prevención de Enfermedades realizan encuestas para comparar los hábitos y actitudes de ciudadanos en varios estados. Para cada uno de los resultados siguientes, proporcione el intervalo de confianza del 95%:
 - a. 17.2% de los encuestados en Colorado son sedentarios (el porcentaje más bajo en Estados Unidos); tamaño de la muestra = 1 500.
 - b. 13.2% de los encuestados en Utah fuman (el porcentaje más bajo en Estados Unidos), tamaño de la muestra = 2500.
 - c. 22.9% de los encuestados en Wisconsin son bebedores consuetudinarios; tamaño de la muestra = 3 500.
22. **Temas importantes.** En una encuesta de ABC/*Washington Post* a 1526 estadounidenses elegidos aleatoriamente se les pidió que listaran los temas más importantes en las elecciones recientes. Educación (79%), economía (74%) y administración del presupuesto (74%) fueron listados como los temas más importantes. El margen de error citado para la encuesta fue de 3 puntos porcentuales. ¿El margen de error es consistente con el tamaño de la muestra?
23. **Drogas en películas.** Un estudio por los investigadores de la Universidad de Stanford para la Oficina Nacional de Política de Control de Drogas y el Departamento de Salud y Servicios Humanos concluyó que 98% de las películas más rentadas involucran drogas, bebida y fumadores. Suponga que el estudio revisó las 400 películas más rentadas:
 - a. Utilice los resultados de esta muestra para estimar la proporción de todas las películas que involucran drogas, bebida o fumadores.
 - b. ¿Cuál es el intervalo de confianza del 95%?
 - c. ¿Usted cree que las 400 películas más rentadas representan una muestra aleatoria? Explique.

- 24. Presión en adolescentes.** Un estudio encargado por el Departamento de Educación de Estados Unidos concluyó que 44% de los adolescentes citan las calificaciones como su fuente principal de presión. El estudio registró respuestas de 1 015 adolescentes. ¿Cuál es el intervalo de confianza del 95%?
- 25. Predicción en elección.** En una muestra aleatoria de 1 600 personas de una gran ciudad, se encontró que 900 apoyan al actual alcalde en la próxima elección. Con base en esta muestra, ¿usted afirmaría que el alcalde ganará con una mayoría de los votos? Explique.
- 26. Predicciones en elección.** En una muestra aleatoria de 1 600 personas de una gran ciudad, se encontró que 900 apoyan al actual alcalde en la próxima elección. Con base en esta muestra, ¿usted afirmaría que el alcalde ganará con una mayoría de los votos? Explique. ¿Qué conclusión hubiese obtenido de una muestra de 250 personas en la cual 130 apoyan al alcalde?
- 27. Sondeo de opinión preelecciones.** Antes de una elección estatal para el senado de Estados Unidos, se realizan tres sondeos de opinión. En la primera encuesta, 780 de 1 500 votantes están a favor del candidato Martínez. En la segunda encuesta, 1 285 de 2 500 votantes están a favor de Martínez. En la tercera, 1 802 de 3 500 votantes están a favor de Martínez. Encuentre intervalos de confianza del 95% para los tres sondeos. Analice las posibilidades de Martínez para ganar con base en estos sondeos.
- 28. Encuesta de desempleo.** La Oficina de Estadísticas del Trabajo estima la tasa de desempleo mensual en Estados Unidos encuestando a 60 000 individuos.
- En un mes, 3.4% de los 60 000 individuos encuestados estaban desempleados. Determine el margen de error para esta estimación. La precisión (al décimo más cercano) es razonable. Explique.
 - Suponga que el número de individuos encuestados se aumentó en un factor de cuatro (a 240 000). ¿En cuánto cambiaría el margen de error?
 - Suponga que el número de encuestados disminuyó en un factor de un cuarto (a 15 000). ¿En cuánto cambiaría el margen de error?
- 29. Encuesta de opinión.** Una encuesta determina que 54% de la población aprueba el trabajo que el presidente está haciendo; la encuesta tiene un margen de error de 4% (suponiendo un grado de 95% de confianza).
- ¿Cuál es el intervalo de confianza del 95% para el porcentaje verdadero de la población que aprueba el desempeño del presidente?
 - ¿Cuál es el tamaño de la muestra para esta encuesta?
- 30. Armas ocultas.** Dos tercios (o 66.6%) de 626 residentes de Colorado encuestados por Talmey-Drake Research & Strategy Inc., dijeron que respaldaban una enmienda pendiente

en la legislatura que estandarizaría leyes sobre el otorgamiento de licencias de armas escondidas a propietarios de armas. La enmienda obligaría a aplicar la ley local para conceder licencia a todo aquel que pueda portar legalmente un arma. El margen de error en la encuesta fue reportado como 4 puntos porcentuales.

- ¿Es el margen de error reportado en consonancia con el tamaño de la muestra para esta estimación?
- ¿Qué tamaño de la muestra sería necesario para dar un margen de error de 2 puntos porcentuales?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 8 en www.aw.com/bbt.

- 31. ¿Quién es el vicepresidente?** Suponga que quiere estimar la proporción de estudiantes en su universidad que pueden identificar de manera correcta al vicepresidente de Estados Unidos. ¿Cuántos estudiantes debe elegir de manera aleatoria para tener una confianza del 95% que su proporción muestral está dentro de 0.1 de la proporción poblacional? Usando ese tamaño de muestra, recolecte sus propios datos muestrales seleccionado aleatoriamente y encuestando a estudiantes de su universidad. Luego utilice los métodos de esta sección para construir un intervalo de confianza del 95%. Escriba una oración que resuma sus resultados.
- 32. Métodos Nielsen.** Visite el sitio web de Nielsen Media Research y reporte los métodos reales usados para estimar proporciones poblacionales e intervalos de confianza en los índices de audiencia Nielsen.
- 33.** Todas las cadenas principales de televisión de manera regular realizan encuestas sobre una variedad de temas. Visite los sitios web de las cadenas principales y reúna los resultados de una encuesta particular que incluya la estimación de una proporción poblacional. Asegúrese de incluir toda la información que esté dada acerca del tamaño de la muestra, margen de error e intervalos de confianza. Incluya cualesquiera detalles respecto al procedimiento real para encuestar.

EN LAS NOTICIAS

- 34. Estimación de proporciones poblacionales.** Encuentre una noticia o reporte en el que una proporción poblacional sea estimada a partir de una muestra. El artículo debe incluir un margen de error y/o un intervalo de confianza. Analice los métodos usados en el estudio y cómo se llegó a las conclusiones.

Ejercicios de repaso del capítulo

1. En un estudio clínico del fármaco Ziac, 3.2% de 221 usuarios experimentaron mareos (con base en información de los laboratorios Lederle).
 - a. Utilice los resultados de esta muestra para construir un intervalo de confianza del 95% que estime la proporción poblacional de usuarios de Ziac que experimentaron mareos.
 - b. Escriba un enunciado que interprete, de manera correcta, el intervalo de confianza encontrado en el inciso a.
 - c. ¿Cuál es el margen de error?
 - d. La proporción muestral de 0.032 se obtuvo de una muestra específica de 221 sujetos. Suponga que muchas muestras diferentes de 221 sujetos se obtienen y que para cada muestra se encontró la proporción de aquellos que experimentaron mareos. ¿Qué sabemos respecto a la forma de la distribución de las proporciones muestrales?
 - e. Si pudiésemos aumentar el tamaño de la muestra de modo que sea mucho mayor a 221, ¿cuál sería el efecto sobre los límites del intervalo de confianza? ¿Estarían más juntos, más alejados o no habría cambio?
 - f. Usando un conjunto diferente de datos, encontramos un intervalo de confianza del 95% $0.400 < p < 0.500$. ¿Qué es incorrecto en esta interpretación: “Existe una probabilidad de 95% que la proporción poblacional caerá entre 0.400 y 0.500”?
2. Queremos estimar la media de la calificación del CI en el examen Stanford-Binet para la población de estudiantes universitarios. Sabemos que para gente elegida aleatoriamente de la población general, la desviación estándar de las calificaciones del CI en el examen Stanford-Binet es 16:
 - a. Usando una desviación estándar de 16, ¿cuántos estudiantes universitarios deben elegirse de manera aleatoria para exámenes del CI, si queremos tener una confianza del 95% de que la media muestral esté dentro de 3 puntos de CI de la media poblacional?
 - b. ¿Cómo se afectaría nuestra estimación de la media del CI de estudiantes universitarios si el tamaño de la muestra fuese mayor de la necesaria? ¿Más pequeña de lo necesario?
 - c. ¿La desviación estándar de las calificaciones del CI para estudiantes universitarios es probable que sea igual a 16, mayor que 16 o menor que 16? Explique. Si utilizamos el valor verdadero en lugar de 16 en el inciso a, ¿cómo se vería afectada la respuesta en el inciso a? ¿Sería la misma, más pequeña o más grande?
3. Se obtuvo una muestra de 50 individuos elegidos aleatoriamente, se cuentan los glóbulos blancos en la sangre de cada individuo. La media es 7.1 y la desviación estándar es 2.5.
 - a. Utilice los resultados de esta muestra para construir un intervalo de confianza del 95% que estime la media poblacional.
 - b. Escriba un enunciado que interprete correctamente el intervalo de confianza encontrado en el inciso a.
 - c. ¿Cuál es el margen de error?
 - d. Si la desviación estándar es 2.5, ¿cuántos individuos deben incluirse, si queremos una confianza del 95% de que la media muestral tenga un error de a lo más 0.25?
4. a. Usted ha sido contratado por Intel para determinar la proporción de propietarios de computadora que planean actualizarlas a un nuevo sistema operativo. Suponiendo que quiere tener una confianza del 95% de que su proporción muestral esté dentro de 0.02 de la proporción poblacional real, ¿cuántas personas debe encuestar?
 - b. Suponga que, al realizar la encuesta descrita en el inciso a, encuentra que la mitad de las personas llamadas rehúsan responder las preguntas de la encuesta, ya que ellos creen que está tratando de venderles algo. Si usted procede a llamar dos veces de modo que el tamaño de su muestra sea suficientemente grande, ¿sus resultados serían buenos? Explique.

Cuestionario del capítulo

1. Si se seleccionan muchas muestras aleatorias diferentes de tamaño 100 de una población de pulsos de mujeres adultas, ¿cuál es la forma de la distribución de las medias muestrales?
2. Si se seleccionan muchas muestras aleatorias diferentes de tamaño 500 de la población de estudiantes universitarios, ¿cuál es la forma de la distribución de las proporciones de mujeres?
3. ¿Qué representa la notación $\hat{p}$?
4. Un artículo de una revista proporciona un intervalo de confianza en el formato 0.60 ± 0.08 . Exprese este intervalo de confianza en el formato $a < p < b$. (Esto es, describa $a < p < b$ usando valores específicos en lugar de a y b).
5. Suponga que queremos estimar la media de la fuerza de sujeción de hombres adultos en Estados Unidos. Si se obtuvo una muestra aleatoria de fuerzas de sujeción, ¿cuál de las siguientes es la mejor estimación de la media poblacional?:
 - a. La mediana de la muestra.
 - b. La media de la muestra.
 - c. La desviación estándar de la muestra.
 - d. El rango de la muestra.
 - e. La proporción muestral.
6. Determine el margen de error correspondiente a este intervalo de confianza del 95%: $0.240 < p < 0.280$.
7. Determine el margen de error correspondiente a este intervalo de confianza del 95%: $240 < \mu < 280$.
8. Cuando un intervalo de confianza del 95% se construye para la media poblacional, una media muestral se encuentra que es 4.60 y el margen de error se encuentra que es 0.15. Identifique el intervalo de confianza del 95%.
9. Cuando se construye un intervalo de confianza del 95% para la proporción poblacional, se encuentra que una proporción muestral es 0.600 y el margen de error es 0.200. Identifique el intervalo de confianza del 95%.
10. Identifique lo que es incorrecto con este intervalo de confianza del 95% para la proporción poblacional: $65.0 < p < 70.0$.

Uso de la tecnología

En la descripción de métodos para la determinación de intervalos de confianza que estimen una media poblacional, este capítulo sólo considera casos en los que se aplica la distribución normal. El capítulo 10 introduce intervalos de confianza para estimaciones de una media poblacional que se encuentran usando la distribución t . Vea los procedimientos al final del capítulo 10 para esos otros métodos.

SPSS

Intervalo de confianza para una media: SPSS no tiene un procedimiento para obtener un intervalo de confianza que estime una media poblacional con desviación estándar poblacional conocida. Vea las instrucciones de tecnología al final del capítulo 10 para un procedimiento que se aplica a situaciones en los que la desviación estándar poblacional no es conocida.

Intervalos de confianza para una proporción: SPSS no tiene un procedimiento para obtener intervalos de confianza para una proporción poblacional.

Excel

Intervalos de confianza que estiman una media poblacional: utilice el complemento Data Desk XL que es un complemento de Excel.

Primero ingrese los datos muestrales en la columna A. Si está usando Excel 2003, dé clic en **DDXL**. Seleccione **Confidence Intervals** (Intervalos de confianza). Debajo de las opciones de tipo de función, seleccione **1 Var z Interval**. Dé clic en el icono del lápiz e ingrese el rango de

datos, tal como A1:A12. Si tiene 12 valores listados en la columna A. Dé clic en **OK**. En el cuadro de diálogo seleccione el nivel de confianza. También introduzca el valor conocido de la desviación estándar poblacional. Dé clic en **Compute Interval** y el intervalo se mostrará.

Intervalos de confianza que estiman una proporción poblacional: utilice el complemento Data Desk XL que es un complemento para Excel.

Primero ingrese el número de éxitos en la celda A1, luego ingrese el número total de ensayos en la celda B1. Si está usando Excel 2003, dé clic en **DDXL**. Si está usando Excel 2007, dé clic en **Adds-in** y luego dé clic en **DDXL**. Seleccione **Confidence Interval**. Seleccione **Summ 1 Var Prop Interval** (que es una forma abreviada en inglés de "intervalo de confianza para una proporción usando datos resumidos para una variable"). Dé clic en el icono del lápiz para "Num successes" e ingrese A1. Dé clic en el icono del lápiz para "Num trials" e ingrese B1. Dé clic en **OK**. En el cuadro de diálogo seleccione el nivel de confianza, luego dé clic en **Compute Interval**.

STATDISK

Seleccione **Analysis** y luego **Confidence Intervals**. Luego seleccione, ya sea **Mean - One Sample** o **Proportion - One Sample**, en el cuadro de diálogo que aparece, primero ingrese el nivel de significancia como un número decimal. Ingrese .95 para un nivel del 95% de confianza. Proceda a ingresar los otros elementos necesarios. Luego dé clic en **Evaluate** y el intervalo de confianza aparecerá.

HABLEMOS DE HISTORIA

¿Dónde inició la estadística?

Los orígenes de muchas disciplinas se pierden en la antigüedad, pero las raíces de la estadística pueden identificarse con alguna certidumbre. La conservación de registro sistemático inició en Londres en 1532 con una colección de datos semanales sobre muertes. Posteriormente en la misma década, en Francia inició una colección de datos oficiales de bautismos, muertes y matrimonios. En 1608, la colección similar a estadísticas vitales inició en Suecia. Canadá realizó el primer censo oficial en 1666.

Por supuesto, la estadística es más que una colección de datos. Si hay un fundador de la estadística, esa persona debe ser alguien que trabajó con datos de un modo inteligente y sistemático y quién usó los datos para llegar a conclusiones que no había sido evidentes anteriormente. Muchos expertos creen que un inglés llamado John Graunt merece el título de fundador de la estadística.

John Graunt nació en Londres en 1620. Como el hijo mayor en una familia numerosa tomó el negocio de su padre como mercero (un comerciante en ropa y alimentos no perecederos). Él pasó la mayor parte de su vida como un prominente ciudadano de Londres, hasta que perdió su casa y sus posesiones en el Incendio de Londres en 1666. Ocho años más tarde murió en la pobreza.

No es claro cómo John Graunt se interesó en los registros semanales de bautismos y obituario —conocidos como registros de mortalidad— que se habían conservado en Londres desde 1563. En el prefacio de su libro *Natural and Political Observations on the Bills of Mortality*, él escribió que otros “hicieron poco uso diferente de ellos” y se sorprendió de “qué provecho del conocimiento del mismo será traído al mundo”. Él debe haber trabajado en sus proyectos estadísticos por muchos años antes de que su libro fuese publicado en 1662.

Graunt trabajó principalmente con registros anuales, los cuales eran resúmenes de fin de año de los registros de mortalidad semanal. La figura 8.14 muestra el registro anual para 1665 (el año de la Gran Plaga). En el tercio superior del registro se muestra el número de sepelios y de bautizos en cada parroquia. El total de sepelios y bautizos se anotó en la parte central del registro, con muertes debidas a la plaga registradas de manera separada. El tercio inferior del registro muestra las muertes debidas a una variedad de causas, con los totales dados para hombres y mujeres.

Graunt estaba consciente de que se habían hecho estimaciones burdas de la población de Londres con propósitos de impuestos, pero él debe haber sido escéptico de hacer una estimación que ponía a la población de Londres en 6 o 7 millones en 1661. Usando los registros anuales, comparando los sepelios y los bautizos, y estimando la densidad de las familias en Londres (con un tamaño promedio de la familia de ocho), llegó a una estimación de la población de 460 000 por tres métodos diferentes —¡muy diferente de 6 o 7 millones! También encontró que la población de Londres estaba aumentando mientras que las poblaciones de los pueblos en el campo estaban disminuyendo, mostrando una tendencia temprana a la urbanización. Él aumentó la conciencia sobre las altas tasas de mortalidad infantil. También refutó una teoría popular de que las plagas llegaban con nuevos reyes.

La contribución más significativa de Graunt puede haber sido su construcción de la primera tabla de vida. Aunque los datos detallados sobre la edad de muerte no estaban disponibles, Graunt sabía que de 100 bebés recién nacidos, “36 de ellos morían antes de cumplir seis años, y que quizá sólo uno sobrevivía a los 76”. Con estos dos datos, él llenó los años intermedios como se muestra en la tabla 8.5, usando métodos que no explicó completamente.



A propósito...

Algunos historiadores aseguran que el libro de Graunt en realidad fue escrito por su amigo de toda la vida y colaborador William Petty. Los argumentos son muy embrollados, pero la mayoría de los estadísticos creen que Graunt escribió su propio libro. De cualquier manera, sabemos que Petty continuó el trabajo de Graunt, publicando ediciones posteriores del libro de Graunt y creando el campo de “aritmética de la política”, que ahora llamamos demografía.

Tabla 8.5 Tabla de vida de Graunt: número de muertes por edad por cada 100 personas

Edad	Muertes
En los primeros seis años	36
Los siguientes 10 años	24
La segunda década	15
La tercera década	9
La cuarta	6
La siguiente	4
La siguiente	3
La siguiente	2
La siguiente	1

Tabla 8.6 Tabla de Graunt de sobrevivientes a varias edades

Edad	Sobrevivientes
A los dieciséis años cumplidos	40
A los veintiséis	25
A los treinta y seis	16
A los cuarenta y seis	10
A los cincuenta y seis	6
A los sesenta y seis	3
A los setenta y seis	1
A los ochenta	0

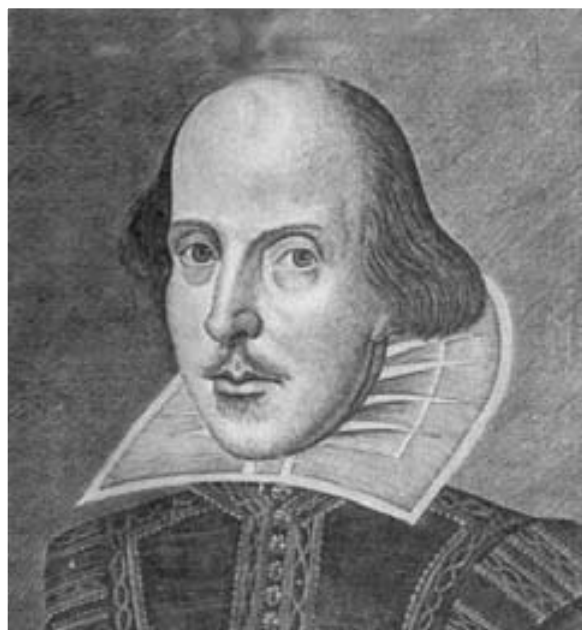
PREGUNTAS PARA DISCUSIÓN

1. ¿Es sorprendente que estimaciones precisas de la población de Londres no estuviesen disponibles en 1660? En esa época, ¿cómo habría sugerido realizar tales estimaciones?
2. ¿Usted considera que los sepelios y bautizos habrían dado cuentas precisas de nacimientos y muertes? ¿Por qué sí o por qué no?
3. Suponiendo que las estimaciones de muertes en la tabla 8.5 sean precisas, ¿las estimaciones de sobrevivientes en la tabla 8.6 son consistentes? Observe que los números en la tabla 8.6 no dan un total de 100; ¿deben darlo? Explique.
4. Con base en las tablas de Graunt, estime la esperanza de vida media en 1660. Explique su razonamiento.

LECTURAS SUGERIDAS

Johnson, N., y Kotz, S. (Editores), *Leading Personalities in the Statistical Sciences*, John Wiley and Sons, 1997.

Sutherland, I., “John Graunt: A Tercentenary Tribute”, *Journal of the Royal Statistical Society*, volumen 126, parte 4, 1963, pp. 537-557.



¿Cuántas palabras conoció Shakespeare?

Imagine que va a una fiesta para estudiantes internacionales. En el transcurso de la noche usted conoce a 12 suecos, 9 chinos, 6 franceses, 4 israelíes, 3 coreanos y 1 iraní. Sabe que existen personas de otras nacionalidades a quienes *no* conoció en la fiesta. Sólo con base en las personas que conoció, ¿es posible estimar el número *total* de nacionalidades representadas en la fiesta —incluyendo a aquellas de las personas que no conoció—? Gracias a las ideas de muestreo, la respuesta es sí.

El problema de la fiesta podría parecer un poco frívolo. Sin embargo, en esencia es la misma pregunta que surgió cuando en la Universidad de Oxford el biólogo marino Charles Paxton se preguntó cuántos “monstruos marinos” (criaturas de más de dos metros de longitud) quedaban por descubrir. En este caso, las nacionalidades en la fiesta corresponden a las especies de monstruos marinos. Usando métodos estadísticos, Paxton fue capaz de estimar que, además de los aproximadamente 220 monstruos marinos ya conocidos, otros 47 esperaban a ser descubiertos.

Métodos similares han sido usados para analizar los trabajos de Shakespeare. Los estadísticos Bradley Efron y Ronald Thisted se preguntaron acerca del número de palabras que realmente conocía Shakespeare, el cual debe haber sido mayor que el número que él utilizó en sus escritos. Aquí, las nacionalidades de la fiesta corresponden a las palabras diferentes en las obras y poemas de Shakespeare. Los datos recolectados en la fiesta pueden considerarse como la *primera muestra*. Para la pregunta de Shakespeare, la primera muestra consiste en todos los trabajos conocidos de Shakespeare —específicamente, los números de palabras que son usadas en estos trabajos una vez, dos veces, tres veces y así sucesivamente—. La tabla 8.7 muestra una parte (muy pequeña) de la primera muestra. Por ejemplo, la tabla dice que en los trabajos de Shakespeare, 14 376 palabras fueron usadas exactamente una vez, 4 343 palabras fueron usadas dos veces, y así sucesivamente. (La tabla completa es mucho más grande y continua para más de 10 ocurrencias).

Dada la tabla completa para la primera muestra, ahora podemos responder una pregunta hipotética. Suponga una segunda, nueva y diferente muestra de trabajos de Shakespeare, del mismo tamaño que la primera muestra, que fuese descubierta. ¿Cuántas palabras esperaríamos encontrar en la segunda muestra que *no* fueron usadas en la primera muestra? ¿Esperaríamos tener menos palabras nuevas en la segunda muestra, ya que en la primera muestra *toda* primera ocurrencia de una palabra es nueva, incluso de una palabra común como “the”; en la segunda muestra, esas palabras comunes ya no son nuevas? Efron y Thisted estimaron que 11 430 palabras que aparecerían en la segunda muestra no aparecían en la primera muestra.

Ellos repitieron este argumento con una tercera muestra, cuarta muestra, quinta muestra, etcétera. Con cada nueva muestra, el número de nuevas palabras disminuía, pero el número total de palabras usadas (entre todas las muestras) aumenta. Efron y Thisted encontraron, eventualmente, que el número de palabras nuevas se aproximaba a cerca de 35 000. Esto significa que además de las 31 534 palabras que Shakespeare conocía y usaba, había aproximadamente 35 000 palabras que él conocía pero no usó. Así, ellos estimaron que Shakespeare conocía aproximadamente 66 500 palabras.

El análisis de Efron y Thisted, que demandó de métodos avanzados, fue hecho en 1976. Más de diez años después, las ideas fueron puestas en práctica cuando fue descubierto un nuevo soneto de un autor desconocido del periodo de Shakespeare. Como antes, el volumen completo de los trabajos de Shakespeare fue considerado la primera muestra. Ahora, la segunda muestra

fue el nuevo soneto con 429 palabras. El mismo método estadístico fue usado para predecir que, si Shakespeare fuese el autor, el nuevo soneto tendría siete palabras nuevas que no aparecerían en los trabajos completos del autor. De hecho, el nuevo soneto tenía nueve palabras nuevas que no aparecían antes. De manera similar, el método predijo que el nuevo soneto debía tener cuatro palabras que fueron usadas exactamente una vez en los trabajos completos. En realidad, había siete palabras que fueron usadas exactamente una vez en los trabajos completos. Y el número de palabras en el soneto que fueron usadas exactamente dos veces en los trabajos completos fue predicho que sería tres, cuando en realidad había cinco de tales palabras. Los autores concluyeron que la concordancia entre las predicciones y las frecuencias reales de las palabras en el soneto era suficientemente buena para atribuir la autoría a Shakespeare.

Estos métodos estadísticos han sido usados para distinguir los trabajos de Shakespeare de otros escritores isabelinos, tales como Marlowe, Donne y Jonson. Los métodos estadísticos han sido usados para fortalecer la creencia que James Madison escribió parte de los Documentos Federales cuya autoría estaba en duda. Incluso los métodos han sido usados para establecer el orden de los trabajos de Platón.

Este ejemplo ilustra la sorprendente aplicabilidad de la estadística a problemas de disciplinas aparentemente no relacionadas. Quizá más importante, demuestra cómo dos disciplinas tan distantes una de otra como la literatura y la biología marina pueden relacionarse por medio de un hilo estadístico común.

Tabla 8.7 Números de palabras en los trabajos completos de Shakespeare usadas de una a diez veces

Ocurrencias	Número de palabras
1	14 376
2	4 343
3	2 292
4	1 463
5	1 043
6	837
7	638
8	519
9	430
10	364

PREGUNTAS PARA DISCUSIÓN

- 1. Comente si cree que la literatura es mejorada por el análisis estadístico. Por ejemplo, ¿su apreciación de Shakespeare mejoró por saber cuántas palabras conocía él?
- 2. ¿Usted cree que el análisis estadístico es útil para identificar autores de “trabajos perdidos”?
- 3. Sugiera otra disciplina, además de biología y literatura, en la cual las ideas descritas en esta sección podrían usarse para estimar una cantidad desconocida.

LECTURAS SUGERIDAS

Efron, B. y Thisted, R., “Estimating the Number of Unknown Species: How many Words Did Shakespeare Know?”, *Biometrika*, volumen 63, número 3, 1976, pp 435-437.

Morin, Richard, “Unconventional Wisdom: Statistics, the King’s Deer, and Monsters of the Sea”, *Washington Post*, 7 de marzo de 1999.

Snell, Laurie, *Chance News* 8.03.



*Es distinguir de una mente educada
ser capaz de considerar una idea
sin aceptarla.*

—Aristóteles

Pruebas de hipótesis

TODOS HACEN AFIRMACIONES. LOS PUBLICISTAS HACEN afirmaciones acerca de sus productos. Las universidades aseguran que sus programas son maravillosos. Los gobiernos afirman que sus programas son efectivos. Los abogados hacen afirmaciones acerca de la culpabilidad o inocencia de un sospechoso. Los diagnósticos médicos son afirmaciones acerca de la presencia o ausencia de una enfermedad. Las compañías farmacéuticas hacen afirmaciones respecto a la eficacia de sus medicamentos. Pero, ¿cómo sabemos si alguna de estas afirmaciones es verdadera? La estadística ofrece una forma para probar muchas afirmaciones, por medio de un conjunto de técnicas poderosas que llevan el nombre de *pruebas de hipótesis*.

OBJETIVOS DE APRENDIZAJE

9.1 Fundamentos de las pruebas de hipótesis

Entender el objetivo de la prueba y la estructura básica de una prueba de hipótesis, incluyendo cómo configurar las hipótesis nula y alternativa, cómo determinar los resultados posibles de dicha prueba, y cómo decidir entre éstos.

9.2 Pruebas de hipótesis para medias poblacionales

Entender e interpretar las pruebas de hipótesis de una y dos colas para afirmaciones realizadas acerca de medias poblacionales, y aprender a reconocer y evitar errores comunes (errores de tipo I y tipo II) en las pruebas de hipótesis.

9.3 Pruebas de hipótesis para proporciones poblacionales

Entender e interpretar pruebas de hipótesis para afirmaciones hechas acerca de proporciones poblacionales.

9.1 Fundamentos de las pruebas de hipótesis

Apropósito...

El producto Gender Choice consistía de instrucciones impresas y un termómetro. Estaba basado en el principio de que el género de un bebé puede determinarse por medio de una cuidadosa sincronización de la concepción. No hubo evidencia de que el producto realmente funcionara y ya no está disponible. Sin embargo, ahora están en el mercado otros productos que parecen tener algún efecto y, por tanto, surgen cuestiones éticas acerca de si y cuándo la elección del género debe ser una opción para los padres.



Una compañía denominada ProCare Industries, Ltd., una vez afirmó que su producto, denominado Gender Choice, podría aumentar la posibilidad de dar nacimiento a una niña. La compañía aseguraba que la posibilidad de un bebé niña aumentaba “hasta 80%”, pero se centraba simplemente en asegurar que la posibilidad de tener una bebé niña es mayor que el aproximadamente 50% o 0.5, esperado bajo circunstancias ordinarias. ¿Cómo podríamos probar si la afirmación de Gender Choice es verdadera?

Una forma podrá ser estudiar una muestra aleatoria de, digamos, 100 bebés nacidos de mujeres que utilizaron el producto Gender Choice. Si el producto no funciona, esperaríamos que alrededor de la mitad de estos bebés sean niñas. Si funciona, esperaríamos que significativamente más de la mitad de los bebés sean niñas. Entonces, la pregunta clave es qué constituye “significativamente más”. Si hubiesen 97 niñas en los 100 nacimientos de la muestra, todos estaríamos de acuerdo en que esto constituye significativamente más de la mitad y que el producto probablemente funcione. Si sólo hay 52 niñas en los 100 nacimientos, quizás estemos de acuerdo que 52 está tan cercano a la mitad que no hay razón para pensar que el producto tuvo algún efecto. Pero ¿consideraríamos, digamos, 64 niñas en la muestra de 100 bebés como “significativamente más” de la mitad y, por tanto, pensaríamos que el producto en realidad podría funcionar?

En estadística respondemos tales preguntas por medio de *pruebas de hipótesis*. Antes de analizar la terminología específica y procedimientos de las pruebas de hipótesis, revisemos el procedimiento que usaríamos para sacar una conclusión acerca de la efectividad de Gender Choice con base en una muestra en la que hay 64 niñas entre 100 bebés nacidos de mujeres que usaron el producto.

- Empezamos suponiendo que Gender Choice *no* funciona; esto es, *no* aumenta el porcentaje de niñas. Si esto es cierto, entonces esperaríamos alrededor de 50% de niñas en la *población* de todos los nacimientos de las mujeres que usaron el producto. Esto es, la proporción de niñas es igual a 0.50.
- Ahora usamos nuestra *muestra* (con 64 niñas en 100 nacimientos) para probar la suposición anterior. Realizamos esta prueba calculando la verosimilitud de obtener una muestra aleatoria de 100 nacimientos en la que 64% (o más) de los bebés son niñas (de una población en la cual la proporción global de niñas es de sólo 50%).
- Si encontramos que una muestra aleatoria de nacimientos es muy probable que tenga 64% (o más) niñas, entonces no tenemos evidencia de que Gender Choice funcione. Sin embargo, si encontramos que una muestra aleatoria de nacimientos es poco probable que tenga al menos 64% de niñas, entonces concluimos que el resultado en la muestra de Gender Choice es probablemente debido a algo distinto del azar, lo que significa que el producto podría ser eficaz. Sin embargo, observe que aunque en este caso no habremos *probado* que el producto es efectivo, todavía podríamos dar otras explicaciones para el resultado (tal como qué sucedió para elegir una muestra tan inusual o que había variables de confusión que no tomamos en cuenta).

UN MOMENTO DE REFLEXIÓN

Suponga que Gender Choice es efectivo y selecciona una nueva muestra aleatoria de 1000 bebés nacidos de mujeres que usaron el producto. Con base en la primera muestra (con 64 niñas en 100 nacimientos), ¿cuántos bebés mujeres esperaríamos en la nueva muestra? Ahora, suponga que Gender Choice *no* es efectivo, ¿cuántas bebés mujeres esperaríamos en una muestra aleatoria de 1000 bebés nacidos de mujeres que usaron el producto?

Formulación de la hipótesis

Los temas clave que quedan en nuestra prueba de la efectividad de Gender Choice son que todavía no hemos descrito cómo calcular la verosimilitud de sacar una muestra aleatoria como aquella con 64% de niñas, ni hemos definido exactamente qué queremos decir cuando pregun-

tamos si tal muestra es “muy probable” o “poco probable” de ser obtenida de manera aleatoria. La mayor parte de este capítulo está dedicado a los procedimientos formales usados para resolver estos temas. El primer paso en su resolución es definir con exactitud qué estamos probando.

Una **hipótesis** es una afirmación acerca de un parámetro poblacional. Por ejemplo, en el caso de Gender Choice, el parámetro poblacional es la proporción de niñas nacidas de todas las mujeres que utilizaron el producto. Entonces, una **prueba de hipótesis** es una prueba para saber si una afirmación particular es respaldada o no por la evidencia disponible.

Definiciones

Una **hipótesis** es una afirmación acerca de un parámetro poblacional, tal como una proporción poblacional (p) o una media poblacional (μ).

Una **prueba de hipótesis** es un procedimiento estándar para probar una afirmación acerca de un parámetro poblacional.

Siempre hay al menos dos hipótesis en cualquier prueba de hipótesis. En el caso de Gender Choice las dos hipótesis, en esencia, son (1) que funciona Gender Choice en aumentar la proporción de la población de niñas del 50% que esperaríamos normalmente y (2) que Gender Choice *no* funciona en aumentar la proporción de la población de niñas. Como ya hemos analizado, el punto inicial para la prueba de hipótesis en realidad es la segunda de esas hipótesis, que *no* aumenta la proporción de niñas. A este punto inicial le llamamos la **hipótesis nula**, o H_0 (se lee “hache cero”). La otra hipótesis (1) se denomina **hipótesis alternativa**, o H_a (se lee “hache a”). Para resumir el ejemplo de Gender Choice:

- La *hipótesis nula* es la afirmación de que Gender Choice *no* funciona, en cuyo caso la proporción poblacional de niñas nacidas de mujeres que usaron el producto debe ser 50% o 0.50. Usando p para representar la proporción poblacional, escribimos esta hipótesis nula como

$$H_0 \text{ (hipótesis nula): } p = 0.50$$

- La *hipótesis alternativa* es la afirmación de que Gender Choice *sí* funciona, en cuyo caso la proporción poblacional de niñas nacidas de mujeres que usaron el producto debe ser *mayor que* 0.50. Escribimos la hipótesis alternativa como

$$H_a \text{ (hipótesis alternativa): } p > 0.50$$

En este capítulo, la hipótesis nula siempre incluirá la condición de igualdad, igual que la hipótesis nula para el ejemplo de Gender Choice es la igualdad $p = 0.50$. (Veremos unos ejemplos de otros tipos de hipótesis nulas en el capítulo 10). Sin embargo, la hipótesis alternativa para el caso Gender Choice ($p > 0.50$) es sólo una de las tres formas generales de hipótesis alternativas que encontraremos en este capítulo:

parámetro poblacional $<$ valor afirmado

parámetro poblacional $>$ valor afirmado

parámetro poblacional $\neq$ valor afirmado

Como veremos en la sección 9.2, estos tres tipos diferentes de hipótesis alternativas necesitan de cálculos ligeramente diferentes en la prueba de hipótesis; por esa razón, es útil darles nombres. La primera forma (“menor que”) conduce a lo que se denomina prueba de hipótesis de **cola izquierda**, ya que necesita probar si el parámetro poblacional está a la *izquierda* (valores inferiores) del valor afirmado. De manera similar, la segunda forma (“mayor que”) lleva a una prueba de hipótesis de **cola derecha**, ya que necesita probar si el parámetro poblacional está a la *derecha* (valores mayores) del valor afirmado. La tercera forma (“diferente”) conduce a una prueba de hipótesis de **dos colas**, puesto que necesita probar si el parámetro de la población está alejado significativamente de *cualquier* lado del valor afirmado.

A propósito...

La palabra *nula* proviene del latín *nullus*, que significa “nada”. Con frecuencia, una hipótesis nula establece que no hay efecto o diferencia especial.

Hipótesis nula e hipótesis alterna

La **hipótesis nula**, o H_0 , es la suposición inicial para una prueba de hipótesis. Para los tipos de pruebas de hipótesis en este capítulo, la hipótesis nula siempre asegura un valor específico para un parámetro poblacional y, por tanto, toma la forma de una igualdad:

H_0 (hipótesis nula): parámetro poblacional = valor afirmado

La **hipótesis alternativa**, o H_a , es una afirmación de que el parámetro poblacional tiene un valor que difiere del valor asegurado en la hipótesis nula. Puede tomar una de las formas siguientes:

(cola izquierda) H_a : parámetro poblacional < valor afirmado

(cola derecha) H_a : parámetro poblacional > valor afirmado

(dos colas) H_a : parámetro poblacional $\neq$ valor afirmado

EJEMPLO 1 Identificación de hipótesis

En cada caso, identifique el parámetro poblacional acerca del cual se hace una afirmación, indique la hipótesis nula y la alterna para una prueba de hipótesis, e indique si la prueba de hipótesis será de cola izquierda, cola derecha o de dos colas.

- a. El fabricante de un nuevo modelo de automóvil híbrido anuncia que el consumo medio de combustible es igual a 62 millas por galón en la autopista. Un grupo de consumidores asegura que la media es menor que 62 millas por galón.
- b. El Departamento de Salud de Ohio asegura que el promedio de estancia en hospitales de Ohio, después del nacimiento de un bebé, es mayor que la media nacional de 2.0 días.
- c. Un biólogo de la fauna que trabaja en la sabana africana afirma que la proporción actual de cebras hembras en la región es diferente de la proporción aceptada de 50%.

Solución

- a. El parámetro poblacional acerca del cual se afirma es una *media* poblacional (μ): la media de millaje de gasolina de todos los automóviles híbridos de este modelo. La hipótesis nula debe indicar que esta media poblacional es *igual* a algún valor específico. Por tanto, reconocemos la hipótesis nula como la afirmación anunciada de que el millaje medio para los automóviles es 62 millas por galón. La hipótesis alternativa es la afirmación del grupo de consumidores que el millaje verdadero de gasolina es *menor* que el anunciado. Para resumir:

H_0 (hipótesis nula): $\mu = 62$ millas por galón

H_a (hipótesis alternativa): $\mu < 62$ millas por galón

Puesto que la hipótesis alternativa tiene una forma “menor que”, la prueba de hipótesis será de cola izquierda.

- b. La población consiste en todas las mujeres de Ohio que recién hayan dado a luz, y el parámetro poblacional es su estancia media (μ) después del alumbramiento en hospitales de Ohio. La hipótesis nula debe ser una igualdad, así que en este caso es la afirmación de que la estancia media en el hospital es igual al promedio nacional de 2.0 días. La hipótesis alternativa es la afirmación del Departamento de Salud que la estancia media en Ohio es *mayor que* este promedio nacional. Para resumir:

H_0 (hipótesis nula): $\mu = 2.0$ días

H_a (hipótesis alternativa): $\mu > 2.0$ días

Puesto que la hipótesis alternativa tiene una forma “mayor que”, la prueba de hipótesis será de cola derecha.

- c. En este caso, la afirmación es acerca de una *proporción* poblacional (p): la proporción de cebras hembras en la población de cebras de la región. La proporción poblacional aceptada

es $p = 0.5$, que se convierte en la hipótesis nula. El biólogo de la fauna afirma que la proporción poblacional real es diferente del valor en la hipótesis nula. Puesto que “diferente” puede ser mayor o menor, la hipótesis alternativa tiene una forma “diferente a”:

$$H_0 \text{ (hipótesis nula): } p = 0.5$$

$$H_a \text{ (hipótesis alternativa): } p \neq 0.5$$

La forma “diferente” para esta hipótesis alternativa lleva a una prueba de dos colas.

Resultados posibles de una prueba de hipótesis

Una prueba de hipótesis siempre inicia suponiendo que la hipótesis nula es verdadera. Luego probamos para ver si la información nos da razón para pensar otra cosa. Como resultado, por lo general sólo hay dos posibles resultados para una prueba de hipótesis, resumidas en el recuadro siguiente.

Los dos resultados posibles de una prueba de hipótesis

Existen dos posibles resultados para una prueba de hipótesis:

1. *Rechazar* la hipótesis nula, H_0 , en cuyo caso tenemos evidencia para respaldar la hipótesis alternativa.
2. *No rechazar* la hipótesis nula, H_0 , en cuyo caso no tenemos evidencia suficiente para respaldar la hipótesis alternativa.

Observe que “aceptar la hipótesis nula” *no* es un resultado posible, ya que la hipótesis nula siempre es la suposición inicial. La prueba de hipótesis podría darnos razón para rechazar esta suposición inicial, pero no puede darnos razón para concluir que la suposición inicial es verdadera.

El que sólo haya dos posibles resultados hace extremadamente importante que la hipótesis nula y la alternativa se elijan de una manera no sesgada. En particular, ambas hipótesis siempre deben formularse *antes* de obtener una muestra de la población para probarlas. De otra manera, los datos de la muestra podrían sesgar la selección de la hipótesis a probar.

UN MOMENTO DE REFLEXIÓN

La idea que una prueba de hipótesis no pueda llevarnos a aceptar la hipótesis nula es un ejemplo del viejo refrán “ausencia de evidencia no es evidencia de ausencia”. Como una ilustración de esta idea, explique por qué sería más fácil, en principio, probar que algún legendario animal (tal como Pie Grande o el Monstruo del Lago Ness) existe, pero es casi imposible probar que *no* existió.

A propósito...

El primer ejemplo registrado de prueba de hipótesis se atribuye al escocés John Arbuthnot (1667-1735). Usando 82 años de datos, él observó que el número anual de bautizos de hombres era consistentemente mayor que el número anual de bautizos de mujeres. Sabiendo que no había sesgo en los bautizos con base en el género, él argumentó que tal patrón regular no podría explicarse por el azar y tenía que ser debido a “la divina providencia”, queriendo decir que los niños debían representar un poco más del 50% de todos los nacimientos. En terminología moderna, él rechazó la hipótesis nula, los datos podrían explicarse sólo por el azar.

EJEMPLO 2 Resultados de una prueba de hipótesis

Para cada uno de los tres casos del ejemplo 1, describa los resultados posibles de una prueba de hipótesis y cómo interpretaríamos estos resultados.

Solución

- a. Recuerde que la hipótesis nula es la afirmación de la publicidad de que la media del millaje para los automóviles nuevos es $\mu = 62$ millas por galón. La hipótesis alternativa es la afirmación del grupo de consumidores que el millaje verdadero es *menor* que el anunciado, o $\mu < 62$ millas por galón. Los resultados posibles son:
 - Rechazar la hipótesis nula de $\mu = 62$ millas por galón, en cuyo caso tenemos evidencia para respaldar la afirmación del grupo de consumidores que el millaje es menor que el anunciado.

- No rechazar la hipótesis nula, en cuyo caso carecemos de evidencia para respaldar la afirmación del grupo de consumidores. Sin embargo, observe que esta opción no implica que la afirmación anunciada sea verdadera.
- b.** La nula es que la estancia media en el hospital es el promedio nacional de 2.0 días. La hipótesis alternativa es la afirmación del Departamento de Salud que la estancia media en Ohio es *mayor que* el promedio nacional. Los resultados posibles son:
- Rechazar la hipótesis nula de $\mu = 2.0$ días, en cuyo caso tenemos evidencia para apoyar la afirmación del Departamento de Salud que la estancia media en Ohio es mayor que el promedio nacional.
 - No rechazar la hipótesis nula, en cuyo caso carecemos de evidencia para apoyar la afirmación del Departamento de Salud. Sin embargo, observe que esta opción no implica que la estancia promedio en Ohio en realidad sea igual a la estancia promedio nacional de 2.0 días.
- c.** La hipótesis nula es que la proporción de cebras hembras es la proporción poblacional aceptada de 50% ($p = 0.5$). La hipótesis alternativa es la afirmación del biólogo que el valor aceptado es incorrecto, esto es, la proporción real de cebras hembras no es 50% (podría ser mayor o menor que 50%). Los posibles resultados son:
- Rechazar la hipótesis nula, en cuyo caso tenemos como evidencia para respaldar la afirmación del biólogo que el valor aceptado es incorrecto.
 - No rechazar la hipótesis nula, en cuyo caso carecemos de evidencia para respaldar la afirmación del biólogo. Sin embargo, observe que esto no implica que el valor aceptado sea correcto.

Cómo sacar una conclusión de una prueba de hipótesis

Regresamos al ejemplo de Gender Choice en el cual ideamos sacar una muestra aleatoria de 100 bebés (nacidos de mujeres que usaron el producto Gender Choice) y encontramos que 64 de estos bebés fueron mujeres. ¿Cómo decidir si esta muestra resultante debe llevarnos a rechazar o no rechazar la hipótesis nula? La respuesta viene de decidir si la muestra resultante era probable o poco probable de que haya ocurrido por azar *si* la hipótesis nula es verdadera.

Recuerde que la hipótesis nula para este caso es que la proporción verdadera de bebés mujeres entre la población de usuarias de Gender Choice es 50%, o $p = 0.50$. Usando la notación introducida en el capítulo 8, la muestra que estamos estudiando tiene un tamaño de $n = 100$ y una proporción muestral $\hat{p} = 0.64$. Entonces, la pregunta precisa es: si la proporción poblacional verdadera es $p = 0.50$ (como lo afirma la hipótesis nula), ¿cuál es la probabilidad de que sólo por azar una muestra de tamaño $n = 100$ tenga una proporción muestral de por lo menos $\hat{p} = 0.64$? Si la probabilidad es baja, entonces es muy poco probable que hubiésemos encontrado tal muestra por azar, por tanto tenemos razones para rechazar la hipótesis nula. Si la probabilidad de observar la muestra resultante es moderada o alta, entonces es muy probable que podríamos haberla obtenido como resultado del azar, por lo que no podemos rechazar la hipótesis nula.

Hay múltiples formas de tomar la decisión acerca de rechazar o no rechazar la hipótesis nula. A continuación revisamos dos opciones muy relacionadas: tomar la decisión con base en la significancia estadística del resultado y tomar la decisión con base en la probabilidad real, o el “valor P ” del resultado de la prueba.

Significancia estadística

Introducimos la idea de significancia estadística en la sección 6.1. Recuerde que si la probabilidad de un resultado particular es 0.05 o menos, decimos que el resultado es estadísticamente significativo al nivel 0.05, si la probabilidad es 0.01 o menor, el resultado es estadísticamente significativo al nivel 0.01. Por tanto, el nivel 0.01 quiere decir que es mayor la significancia que la del nivel 0.05. El recuadro siguiente resume cómo podemos aplicar estas ideas de manera directa a las pruebas de hipótesis.

La verdad probable depende de argumentos estadísticos para valorar cuál de las diferentes posibilidades es más probable que sea verdadera. Cuando algo debe ser probado más allá de una duda razonable, ¿eso qué significa? ¿Qué nivel de duda es aceptable? ¿Uno en veinte? ¿Uno en un billón?

—K. C. Cole

Decisiones en pruebas de hipótesis basadas en niveles de significancia estadística

Decidimos el resultado de una prueba de hipótesis comparando el resultado muestral real (media o proporción) con el resultado esperado *si* la hipótesis nula es verdadera. Debemos elegir un nivel de significancia para la decisión.

- Si la posibilidad de un resultado muestral al menos tan extremo como el resultado observado es menor de 1 en 100 (o 0.01), entonces la prueba es estadísticamente significativa al nivel 0.01 y ofrece una fuerte evidencia para rechazar la hipótesis nula.
- Si la posibilidad de un resultado muestral al menos tan extremo como el resultado observado es menor de 1 en 20 (o 0.05), entonces la prueba es estadísticamente significativa al nivel 0.05 y ofrece una evidencia moderada para rechazar la hipótesis nula.
- Si la posibilidad de un resultado muestral al menos tan extremo como el resultado observado es mayor que el nivel de significancia elegido (0.05 o 0.01), entonces no rechazamos la hipótesis nula.

EJEMPLO 3 Significancia estadística en pruebas de hipótesis

Considere el ejemplo de Gender Choice en el que el tamaño de la muestra es $n = 100$ y la proporción muestral $\hat{p} = 0.64$. Usando técnicas que analizaremos posteriormente en este capítulo, es posible calcular la probabilidad de elegir de manera aleatoria tal muestra (o una muestra más extrema con $\hat{p} > 0.64$) bajo la suposición que la hipótesis nula ($p = 0.50$) es verdadera; el resultado es que la probabilidad es 0.0026. Con base en este resultado, ¿debemos rechazar o no la hipótesis nula?

Solución La probabilidad de 0.0026 significa que si la hipótesis nula es verdadera (queriendo decir que la proporción poblacional verdadera es 50%), la probabilidad de extraer una muestra aleatoria con una proporción muestral de por lo menos $\hat{p} = 0.64$ es sólo menor que 3 en 1000. Puesto que esta probabilidad es menor que 1 en 100, este resultado es estadísticamente significativo al nivel 0.01. Por tanto, nos da una buena razón para rechazar la hipótesis nula, lo que significa que proporciona respaldo para la hipótesis alternativa que el producto Gender Choice aumenta la proporción de niñas a más de 50%.

Valores p

En el ejemplo 3 anterior concluimos que el resultado muestral nos da razón para rechazar la hipótesis nula, ya que el resultado es estadísticamente significativo al nivel 0.01. De hecho, el resultado es aún mejor que eso: la probabilidad calculada de 0.0026 es alrededor de un cuarto de 0.01. Por tanto, la probabilidad precisa nos da aún más información que sólo establecer un nivel de significancia estadística.

Los resúmenes publicados o reportes de noticias de pruebas de hipótesis con frecuencia sólo indican el nivel de significancia, pero la determinación de ese nivel siempre requiere que primero calculemos una probabilidad precisa. Esta probabilidad se denomina **valor P** (abreviación de *valor de probabilidad*); observe la letra mayúscula P , usada para evitar la confusión con la letra minúscula p que se establece para la proporción poblacional. En otras palabras, para el caso del ejemplo 3, en el que la probabilidad de extraer una muestra con una proporción muestral de al menos 64% es 0.0026, decimos que el valor P para la prueba de hipótesis es 0.0026. Analizaremos el cálculo real del valor P en las secciones 9.2 y 9.3; aquí, sólo nos concentraremos en su interpretación.

Decisiones de pruebas de hipótesis con base en valores P

El **valor P** (valor de probabilidad) para una prueba de hipótesis de una afirmación acerca de un parámetro poblacional es la probabilidad de seleccionar una muestra al menos tan extrema como la muestra observada, suponiendo que la hipótesis nula es verdadera:

- Un valor P pequeño (tal como menor o igual a 0.05) indica que la muestra resultante es poco probable y por tanto proporciona razón para rechazar la hipótesis nula.
- Un valor P grande (mayor que 0.05) indica que la muestra resultante podría fácilmente haber ocurrido por azar, por lo que no podemos rechazar la hipótesis nula.

EJEMPLO 4 ¿Moneda justa?

Usted sospecha que una moneda podría estar sesgada hacia la cruz con mayor frecuencia que cara, y decide probar esta sospecha lanzando la moneda 100 veces. El resultado es que obtiene 40 caras (y 60 cruces). Un cálculo (que no se muestra aquí) indica que la probabilidad de obtener 40 o menos caras en 100 lanzamientos con una moneda no sesgada es 0.0228. Determine el valor P y el nivel de significancia estadística para su resultado. ¿Debe concluir que la moneda está sesgada en contra de las caras?

Solución La hipótesis nula es que la moneda es no sesgada, en cuyo caso la proporción de caras debe ser alrededor de 50% ($H_0: p = 0.50$). La hipótesis alternativa es su sospecha de que la moneda está sesgada en contra de la cara, en cuyo caso la proporción de caras sería menor que 50% ($H_a: p < 0.50$). Sus 100 lanzamientos de la moneda representan una muestra aleatoria de tamaño $n = 100$, y el resultado de 40 caras es la proporción muestral ($\hat{p} = 0.40$) para la prueba de hipótesis. El valor P para la prueba es la probabilidad de obtener una proporción muestral al menos tan extrema como la que usted encontró ($\hat{p} \leq 0.40$), *suponiendo* que la moneda es no sesgada y la proporción poblacional de caras es 0.5. La probabilidad dada para esta ocurrencia es 0.0228, que es el valor P para la prueba. Puesto que este valor P es menor que 0.05, el resultado es estadísticamente significativo al nivel 0.05. Puesto que no es menor que 0.01, el resultado no es estadísticamente significativo al nivel 0.01. La significancia estadística al nivel 0.05 nos da una razón moderada para rechazar la hipótesis nula y concluir que la moneda está sesgada en contra de las caras.

Reunir todo

Ahora hemos cubierto todas las ideas básicas de la prueba de hipótesis, excepto los cálculos reales requeridos. Analizaremos estos cálculos para las pruebas de hipótesis de las medias poblacionales en la sección 9.2 y para las proporciones poblacionales en la sección 9.3. El recuadro siguiente resume los pasos que se realizan en una prueba de hipótesis.

El proceso de una prueba de hipótesis

- Paso 1. Formular las hipótesis nula y la alternativa, cada una de las cuales debe hacer una afirmación acerca de un parámetro *poblacional*, tal como una media poblacional (μ) o una proporción poblacional (p); asegúrese de hacer esto antes de sacar una muestra o recolectar datos. Con base en la forma de la hipótesis alternativa, decida si necesitará una prueba de cola izquierda, derecha o de dos colas.
- Paso 2. Saque una muestra de la población y mida los estadísticos muestrales, incluyendo el tamaño de la muestra (n) y el estadístico muestral relevante, tal como la media muestral ($\bar{x}$) o la proporción muestral ($\hat{p}$).
- Paso 3. Determine la verosimilitud de observar un estadístico muestral (media o proporción) al menos tan extremo como el que encontró *bajo la suposición que la hipótesis nula es verdadera*. La probabilidad precisa de tal observación es el valor P (valor de probabilidad) para su muestra.
- Paso 4. Decida si rechaza o no rechaza la hipótesis nula en su nivel de significancia elegido (por lo común 0.05 o 0.01, aunque en algunas ocasiones, otros niveles de significancia son usados).

Nuevamente, asegúrese de evitar confusión entre los tres diferentes usos de la letra p :

- Una letra p minúscula representa una *proporción poblacional*, esto es, la proporción verdadera en toda la población.
- Una letra minúscula $\hat{p}$ (“ p -gorro”) representa una *proporción muestral*, esto es, la proporción encontrada en una muestra extraída de la población.
- Una P mayúscula indica la probabilidad en el valor P .

EJEMPLO 5 Millaje medio de renta de automóviles

En Estados Unidos, el automóvil promedio se maneja alrededor de 12 000 millas cada año. El propietario de una gran compañía de renta de automóviles sospecha que para su flotilla, la distancia media es mayor que 12 000 millas por año. Él selecciona una muestra aleatoria de $n = 225$ automóviles de su flotilla y determina que el millaje medio anual para esta muestra es $\bar{x} = 12\,375$ millas. Un cálculo muestra que si usted supone que la media de la flotilla es la media nacional de 12 000 millas, entonces la probabilidad de seleccionar una muestra de 225 automóviles con un millaje medio de al menos 12 375 millas es 0.01. Con base en estos datos, describa el proceso de llevar a cabo una prueba de hipótesis y saque una conclusión.

Solución Seguimos los cuatro pasos listados en el recuadro anterior:

Paso 1. El parámetro poblacional de interés es una *media poblacional* (μ) —la media de millaje anual para la población de todos los automóviles en la flotilla de automóviles en renta—. La hipótesis nula es que la media poblacional es el promedio nacional de 12 000 millas anuales por automóvil. La hipótesis alternativa es la afirmación del propietario que la media poblacional de su flotilla es mayor que el promedio nacional. Esto es,

$$H_0: \mu = 12\,000 \text{ millas}$$

$$H_a: \mu > 12\,000 \text{ millas}$$

Puesto que la hipótesis alternativa tiene una forma “mayor que”, la prueba de hipótesis será de cola derecha.

Paso 2. Este paso pide que seleccionemos la muestra y midamos los estadísticos muestrales; aquí, se nos dan el tamaño de la muestra de $n = 225$ y la media muestral de $\bar{x} = 12\,375$ millas.

Paso 3. Este paso nos pide determinar la verosimilitud que, bajo la suposición que el promedio de la flotilla en realidad sea 12 000 millas (hipótesis nula), seleccionásemos por azar una muestra con una media de por lo menos 12 375 millas. Aquí no necesitamos hacer los cálculos, ya que nos dieron la probabilidad de 0.01. (El método de cálculo se da en la sección 9.2). Esta probabilidad es el valor P ; nos dice que si la hipótesis nula fuese verdadera, habría sólo una probabilidad de 0.01 de seleccionar una muestra tan extrema como la observada.

Paso 4. El valor P de 0.01 nos dice que el resultado es significativo al nivel 0.01. Por tanto, rechazamos la hipótesis nula y concluimos que la prueba proporciona fuerte apoyo para la hipótesis alternativa, implicando que la media anual de millaje para la flotilla de renta de automóviles es mayor que el promedio nacional de 12 000 millas.

Una analogía legal de prueba de hipótesis

Una analogía legal podría ayudar a clarificar la idea de prueba de hipótesis. En las cortes legales estadounidenses, el principio fundamental es que un acusado se presume inocente hasta que se pruebe que es culpable. Puesto que la suposición inicial es inocente, esto representa la hipótesis nula:

$$H_0: \text{El acusado es inocente.}$$

$$H_a: \text{El acusado es culpable.}$$

El trabajo del fiscal es presentar evidencia tan convincente que el jurado se persuada de rechazar la hipótesis nula y encuentre al acusado culpable. Si el fiscal no construye un caso suficientemente convincente, entonces el jurado no rechazará H_0 y el acusado se encontrará “no culpable”. Observe que encontrar a una persona inocente (aceptando H_0) no es una opción: un veredicto de no acusado significa que la evidencia no es suficiente para establecer culpabilidad, pero no prueba inocencia.



UN MOMENTO DE REFLEXIÓN

Considere dos situaciones. En una, usted es jurado en un caso en el que el acusado podría ser multado con un máximo de \$2000. En el otro, usted es un jurado en un caso en el cual el acusado podría recibir la pena de muerte. Compare los niveles de significancia que usaría en las dos situaciones. En cada situación, ¿cuáles son las consecuencias de rechazar de manera errónea la hipótesis nula?

Sección 9.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- 1. Prueba de hipótesis.** ¿Qué es una prueba de hipótesis?
 - 2. Hipótesis.** ¿Qué es la hipótesis nula? ¿Qué notación se utiliza para la hipótesis nula? ¿Qué es una hipótesis alternativa? ¿Qué notación se utiliza para una hipótesis alternativa?
 - 3. Hipótesis alternativa.** Un investigador quiere probar una afirmación hecha acerca de la temperatura corporal de adultos saludables; la afirmación se refiere al valor de 98.6°F. Identifique las tres posibles expresiones diferentes que podrían usarse para la hipótesis alternativa.
 - 4. Valor P .** ¿Qué es el valor P para una prueba de hipótesis?
- ¿Tiene sentido?** Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.
- 5. Reporte en medios.** La revista *Newport Chronicle* incluye un artículo con el enunciado que “el estudio probó que el porcentaje de aspirantes a trabajos que no pasaban una prueba de drogas es igual a 6%”.
 - 6. Valor P .** Una investigadora está convencida que ella puede demostrar que un medicamento nuevo es efectivo en disminuir el colesterol LDL. Ella afirma que el valor P de 0.001 respalda su afirmación de un nivel medio más bajo de colesterol LDL.
 - 7. Hipótesis nula.** Al probar una afirmación de que el nivel medio de colesterol LDL es menor que 130 mg/dL, el investigador establece la hipótesis nula como $\mu < 130$ mg/dL.
 - 8. Hipótesis nula.** Después de realizar una prueba de hipótesis, un investigador forma una conclusión inicial de que hay evidencia muestral estadística para respaldar la hipótesis nula.
 - 9. Tratamiento con medicamento.** En una prueba de la afirmación de que, entre los pacientes tratados con Ziac, la proporción que experimentó mareos es menor que 0.06, la hipótesis alternativa es $p \geq 0.06$.

- 10. Media muestral.** Un estudio está diseñado para determinar la proporción de hombres que pesan más de 200 libras, por lo que debe encontrarse el peso media muestral.
- 11. Valor P .** En una prueba clínica de un medicamento nuevo, el coordinador del proyecto está feliz de saber que la efectividad del medicamento está descrita con un valor P de 0.001.
- 12. Pulso.** En una prueba de la afirmación de que quien hace ejercicio tiene un pulso medio menor que 74 latidos por minuto, la hipótesis nula es $\mu = 74$ y la hipótesis alternativa es $\mu < 74$.

Conceptos y aplicaciones

- 13. ¿Qué es la significancia?** Suponga que usted viaja diario a la universidad y llena su tanque de gasolina cada viernes por la tarde con 7 galones de gasolina.
 - a. Una semana, después de un manejo y viaje normal, usted encuentra que su automóvil necesita 8 galones. ¿Consideraría este aumento estadísticamente significativo o lo atribuiría a fluctuaciones aleatorias?
 - b. Después de otra semana de manejar y viajar normalmente, encuentra que su automóvil necesita 12 galones. Este aumento, ¿lo consideraría estadísticamente significativo o lo atribuiría a variaciones aleatorias?
 - c. ¿Qué tanto aumento necesitaría ver antes de que lo considere estadísticamente significativo y buscar una explicación? Explique.
- 14. ¿Qué es significativo?** Suponga que sus cargos de llamadas telefónicas de larga distancia son de \$7 mensuales.
 - a. Después de un mes de uso normal, el monto de las llamadas telefónicas es \$8.05. ¿Consideraría este cambio estadísticamente significativo y buscaría una explicación, o atribuiría este tipo de variación al azar?
 - b. Después de un mes de uso normal, el monto de las llamadas telefónicas es \$40. ¿Consideraría este cambio estadísticamente significativo y buscaría una explicación, o atribuiría este tipo de variación al azar?
 - c. ¿Alrededor de cuánto aumento necesitaría ver antes de que lo considere estadísticamente significativo y buscar una explicación? Explique.

Formulación de hipótesis. En los ejercicios 15 al 22 formule la hipótesis nula y la alternativa para una prueba de hipótesis. Indique en términos claros las dos posibles conclusiones a las que se puede llegar con la afirmación dada.

15. **Contenidos de refresco.** Un consumidor se queja de que las botellas de cola suministradas por un supermercado en realidad contienen menos de un litro de refresco.
16. **Calificaciones del SAT.** Un director de preparatoria afirma que, en su escuela, la calificación media del SAT de los alumnos de último año es menor que el promedio nacional de 1 518.
17. **Selección de género.** El director de Operaciones de un centro médico afirma que los tratamientos pueden aumentar la probabilidad de que un bebé será mujer por lo que la proporción de niñas es mayor que 0.7.
18. **Control de calidad.** El gerente de control de calidad en una compañía fabricante de caminadoras asegura que la proporción de caminadoras defectuosas es menor que 0.02.
19. **Máquinas expendedoras.** Una representante de ventas afirma que sus máquinas expendedoras de café lo expenden de modo que la cantidad media suministrada es igual a 10 onzas.
20. **Aspirina.** La Administración de Alimentos y Drogas afirma que una compañía farmacéutica está produciendo tabletas de aspirina con una cantidad media de aspirina que es menor a 350 miligramos.
21. **Holocausto.** Una maestra de preparatoria afirma que la mayoría de sus estudiantes no conocen el significado de *Holocausto*.
22. **Fumar.** Un educador asegura que menos del 20% de graduados universitarios fuma.

Valor P y nacimientos. Suponga que los nacimientos de hombre y de mujer son igualmente probables. La tabla siguiente muestra las probabilidades de diferentes números de hombres en una muestra aleatoria de 100 nacimientos. Utilice esta información para los ejercicios 23 al 28, y suponga que estamos probando un sesgo en contra de hombres.

Número de hombres en 100	Probabilidad
35 o menos	0.002
40 o menos	0.028
45 o menos	0.184
48 o menos	0.382

23. Una muestra aleatoria de 100 nacimientos tiene 48 hombres. ¿Este resultado es significativo al nivel 0.05? ¿Cuál es el valor P de este resultado?

24. Una muestra aleatoria de 100 nacimientos tiene 45 bebés hombres. ¿Este resultado es significativo al nivel 0.01? ¿Cuál es el valor P de este resultado?
25. Una muestra aleatoria de 100 nacimientos tiene 40 bebés hombres. ¿Este resultado es significativo al nivel 0.01? ¿Cuál es el valor P de este resultado?
26. Una muestra aleatoria de 100 nacimientos tiene 40 bebés hombres. ¿Este resultado es significativo al nivel 0.05? ¿Cuál es el valor P de este resultado?
27. Una muestra aleatoria de 100 nacimientos tiene 35 bebés hombres. ¿Este resultado es significativo al nivel 0.01? ¿Cuál es el valor P de este resultado?
28. Una muestra aleatoria de 100 nacimientos tiene 32 bebés hombres. ¿Este resultado es significativo al nivel 0.01? ¿Cuál es el valor P de este resultado?



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 9 en www.aw.com/bbt.

29. **Revistas profesionales.** Muchas revistas profesionales, como *Journal of the American Medical Association*, contienen artículos que incluyen información acerca de pruebas de hipótesis formales. Encuentre un artículo de éstos e identifique la hipótesis nula y la hipótesis alternativa. En términos sencillos, indique el objetivo de la prueba de hipótesis y la conclusión a la que se llegó.
30. **Actividad con una moneda.** Seleccione una moneda de 5 pesos y pruebe la afirmación de que está a favor de sol cuando se lanza. Indique la hipótesis nula y la alternativa, luego lance la moneda 100 veces. Aplicando sólo sentido común, ¿cuál es su conclusión acerca de la afirmación de que está a favor de que aparezca sol?

EN LAS NOTICIAS

31. **Prueba de hipótesis en las noticias.** Encuentre una noticia o un reporte de investigación que describa (quizá no de manera explícita) una prueba de hipótesis para una media o una proporción poblacionales. Adjunte el artículo y resuma el método utilizado.

9.2 Pruebas de hipótesis para medias poblacionales

En la sección 9.1 hicimos un bosquejo del resumen básico del proceso de pruebas de hipótesis. Recuerde que para todos los casos, el objetivo final de la prueba es tomar una decisión acerca de si se rechaza o no se rechaza la hipótesis nula, que es la suposición inicial para la prueba. En esta sección describimos los cálculos usados para tomar esa decisión para las pruebas de hipótesis con medias poblacionales. (En la sección siguiente tratamos las proporciones muestrales). Primero investigamos el procedimiento para pruebas de una cola (cola izquierda o cola derecha) y luego analizamos las pequeñas diferencias en el procedimiento necesario para pruebas de dos colas.

Pruebas de hipótesis de una cola

Considere la siguiente situación hipotética. Columbia College anuncia que el salario medio de sus graduados es de \$39 000. El Comité para la Verdad en la Publicidad, una organización independiente, sospecha que está exagerando y decide llevar a cabo una prueba de hipótesis para buscar evidencia que respalde su sospecha.

Observa que el parámetro de interés es el salario medio inicial de la población de todos los graduados de Columbia College, por lo que la prueba de hipótesis tendrá que ver con una *media poblacional* (μ). La hipótesis nula es la afirmación del comité de que la universidad está exagerando el salario medio inicial en cuyo caso el salario medio inicial es *menor que* \$39 000. En otras palabras, la hipótesis nula y la alternativa son

$$H_0: \mu = \$39\,000$$

$$H_a: \mu < \$39\,000$$

Puesto que la hipótesis alternativa tiene una forma “menor que”, estamos tratando con una prueba de hipótesis de cola izquierda. El procedimiento general para pruebas de cola derecha o izquierda es el mismo, por lo que consideraremos ambas como pruebas de *una cola*.

Habiendo construido la hipótesis, el Comité para la Verdad en la Publicidad selecciona una muestra aleatoria de 100 graduados recientes del colegio. El salario medio de los graduados en la muestra resulta ser \$37 000. En otras palabras, el tamaño de la muestra es $n = 100$ y la media muestral es $\bar{x} = \$37\,000$.

Si ahora revisa los cuatro pasos en las pruebas de hipótesis en la página 376, verá que los primeros dos pasos ya están completos: el salario medio inicial ha sido identificado como el parámetro poblacional de interés, las hipótesis nula y alternativa se han establecido, y la muestra se ha sacado y medido para determinar el tamaño de la muestra y la media muestral. Por tanto, estamos preparados para los pasos 3 y 4, en los que suponemos que la hipótesis nula es verdadera y entonces determinar si el estadístico muestral nos da la razón para rechazar esta suposición.

La distribución muestral

El paso 3 del proceso de pruebas de hipótesis es encontrar la verosimilitud de observar una media muestral tan extrema como la encontrada, bajo la suposición que la hipótesis nula es verdadera. Para el ejemplo de Columbia College, la pregunta se convierte en: ¿cuán probable es que seleccionemos una muestra (de tamaño $n = 100$) con una media de \$37 000 o menos cuando la media para la población completa es \$39 000? Para responder esta pregunta necesitamos una observación crítica con base en nuestro trabajo con distribuciones de muestreo del capítulo 8: *la media muestral observada* ($\bar{x} = \$37\,000$) *es sólo un punto en una distribución de medias muestrales*. Además, suponiendo que la hipótesis nula sea verdadera, esta distribución de medias muestrales es aproximadamente normal con una media de $\mu = \$39\,000$.

A propósito...

A partir de 2006, datos del Centro Nacional para Estadísticas Educativas indican que la mediana del salario inicial para *todos* los graduados de universidad con un grado de licenciatura es alrededor de \$35 000. La mediana está dada en lugar de la media, porque la distribución del salario tiende a ser muy sesgada con grandes variaciones por área. Los graduados en ingeniería y los de administración por lo general tienen los salarios iniciales más altos.

Para entender esta observación, imagine que en lugar de sacar sólo una muestra de tamaño $n = 100$, el Comité para la Verdad en la Publicidad saca muchas muestras de ese tamaño. Cada muestra tendría una única media muestral, $\bar{x}$, por lo que podríamos hacer una gráfica de la distribución de todas estas medias muestrales. Como se estudió en la sección 8.1, para una muestra razonablemente grande tal como $n = 100$, la distribución de muestreo resultante será aproximadamente normal con una media que es igual a la media de la población. Puesto que la hipótesis nula afirma que la media poblacional es $\mu = \$39\,000$, la distribución muestral tendrá un pico en este valor, si la hipótesis nula es correcta.

La figura 9.1 muestra gráficamente la idea. La curva en gris representa la distribución muestral suponiendo que la hipótesis nula es verdadera. Recuerde que esta curva es la que esperaríamos encontrar si graficásemos la distribución de las medias de *muchas* muestras. Cuando sólo tenemos una media muestral (tal como la media muestral $\bar{x} = \$37\,000$ para el ejemplo de Columbia College), sólo representa un punto en esta curva. Si este punto está cercano al pico de la curva, nos indica que la media muestral está cerca de la media poblacional esperada con la hipótesis nula. En ese caso, la probabilidad de encontrar tal media muestral no es pequeña y no hay razón para rechazar la hipótesis nula. En contraste, si la media muestral está lejos de la media poblacional afirmada por la hipótesis nula, entonces la probabilidad de encontrar tal media muestral es pequeña *si* la hipótesis nula es verdadera. Por tanto, concluiríamos que la media poblacional verdadera probablemente *no* es la que la hipótesis nula afirma, en cuyo caso tenemos razón para rechazar la hipótesis nula.

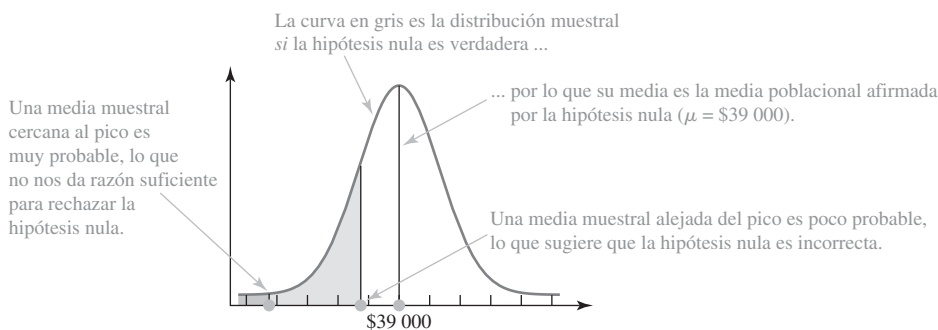


Figura 9.1 Interpretación gráfica de la prueba de hipótesis para el ejemplo de Columbia College. Si la hipótesis nula es verdadera, entonces la media población es $\mu = \$39\,000$. En ese caso, si tomamos muchas muestras de la población, la distribución de las medias muestrales sería aproximadamente normal con una media de $\$39\,000$. Bajo esta suposición, la distancia de una sola media muestral a la media de la población nos permite decidir rechazar o no rechazar la hipótesis nula.

Determinación de la puntuación estándar

Como se ilustró en la figura 9.1, la decisión de rechazar la hipótesis nula depende de si la media muestral está “cerca” o “lejos” de la media poblacional afirmada por la hipótesis nula. Puesto que la distribución muestral es una distribución (casi) normal, el *error estándar* de la media muestral representa una medida cuantitativa de la distancia entre la media muestral y la media poblacional afirmada. Recuerde de la sección 5.2 que la puntuación estándar (puntuación z) de un valor en una distribución normal es el número de desviaciones estándar que está por arriba o por abajo de la media de la distribución. Del teorema del límite central (sección 5.3), la desviación estándar de una distribución de medias muestrales es $\sigma/\sqrt{n}$, donde σ es la desviación estándar poblacional y n es el tamaño de la muestra. Reuniendo estas ideas, encontramos la fórmula siguiente para la puntuación estándar de una media muestral $\bar{x}$:

$$z = \frac{\text{media muestral} - \text{media poblacional}}{\text{desviación estándar de la distribución muestral}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

El único problema restante es que por lo general no conocemos la desviación estándar poblacional, σ . Por ahora, supongamos que podemos aproximar la desviación estándar de la población con la desviación estándar muestral, s , y que $s = \$6\,150$ para los 100 salarios en

la muestra. (En la sección 10.1 analizaremos una mejor forma de proceder cuando σ no es conocida). En ese caso, tomamos $\sigma = \$6150$ y la desviación estándar para la distribución de medias muestrales es

$$\frac{\sigma}{\sqrt{n}} = \frac{\$6150}{\sqrt{100}} = \$615$$

Usando este valor en la ecuación anterior nos dice que la puntuación estándar para la media muestral de $\bar{x} = \$37000$ en una distribución muestral con una media poblacional de $\mu = \$39000$ es

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\$37000 - \$39000}{\$615} = -3.25$$

En otras palabras, la media muestral de $\bar{x} = \$37000$ está a 3.25 desviaciones estándar por debajo de la media de la distribución muestral.

La figura 9.2 muestra que la puntuación estándar de -3.25 ubica el resultado muestral lejos en la cola izquierda de la distribución, indicando que sería raro un resultado muestral si la media poblacional en realidad fuesen los $\$39000$ afirmados por la hipótesis nula. Esto sugiere que la hipótesis nula es incorrecta y debemos rechazarla. Sin embargo, en lugar de rechazar la hipótesis nula sólo por esta razón visual, revisemos un procedimiento formal para analizar la puntuación estándar y tomar la decisión de la prueba de hipótesis.

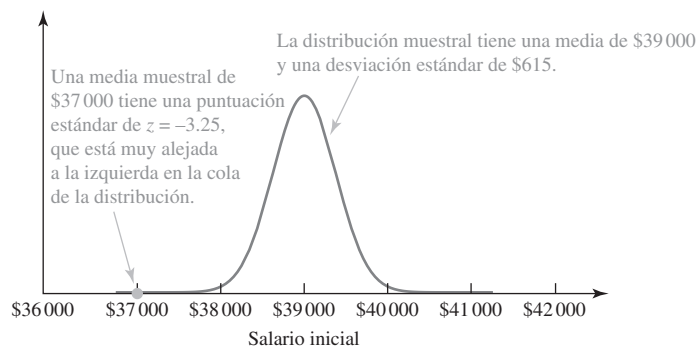


Figura 9.2 Suponiendo que la hipótesis nula es verdadera, la distribución de las medias muestrales para el ejemplo de Columbia College tiene una media de $\$39000$ y una desviación estándar de $\$615$. En ese caso, una media muestral de $\bar{x} = \$37000$ tiene una puntuación estándar de -3.25 .

NOTA TÉCNICA

La puntuación estándar para la media muestral en ocasiones se llama *estadístico de prueba*.

Cálculo de la puntuación estándar para la media muestral en una prueba de hipótesis

Cuando sacamos una media muestral para una prueba de hipótesis, podemos considerarla como una de las muchas muestras posibles en la distribución muestral. Dado el tamaño de la muestra (n), la media muestral ($\bar{x}$), la desviación estándar poblacional (σ) y la media población afirmada (μ), hacemos los cálculos siguientes:

$$\text{desviación estándar para la distribución de medias muestrales} = \frac{\sigma}{\sqrt{n}}$$

$$\text{puntuación estándar para la media muestral, } z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Observación: en realidad, es raro que conozcamos la desviación estándar poblacional σ ; vea la sección 10.1 acerca de cómo tratar con tales casos.

Valores críticos para significancia estadística

Recuerde que la prueba de hipótesis es significativa al nivel 0.05 si la probabilidad de encontrar un resultado tan extremo como el observado realmente es 0.05 o menos (suponiendo que la hipótesis nula es verdadera). Entonces para una prueba de cola izquierda buscamos una puntuación estándar que esté en o debajo del quinto percentil de la distribución muestral. De la tabla 5.1 (página 211), el quinto percentil tiene una puntuación estándar entre $z = -1.6$ y $z = -1.7$; las tablas más precisas de puntuación estándar en el apéndice A muestran que el quinto percentil tiene una puntuación estándar de $z = -1.645$. Por tanto, una hipótesis de cola izquierda es significativa al nivel 0.05 si la puntuación estándar de la media muestral es menor o igual a $z = -1.645$. Esta puntuación estándar representa el **valor crítico** para significancia al nivel 0.05 en una prueba de hipótesis de cola izquierda.

Un argumento similar se aplica a pruebas de cola derecha con las hipótesis alternativas de la forma $H_a: \mu > \text{valor afirmado}$. En tales casos, la significancia al nivel 0.05 requiere que la media muestral esté en o por arriba del percentil 95o., lo que necesita una puntuación estándar mayor o igual a $z = 1.645$. La figura 9.3 ilustra estos valores críticos para las pruebas de cola izquierda y de cola derecha. Para encontrar los valores críticos para significancia al nivel 0.01, buscamos las puntuaciones estándar de los percentiles primero y 99o. (en lugar del 5o. y 95o.). El apéndice A muestra que éstos son -2.33 y 2.33 , respectivamente.

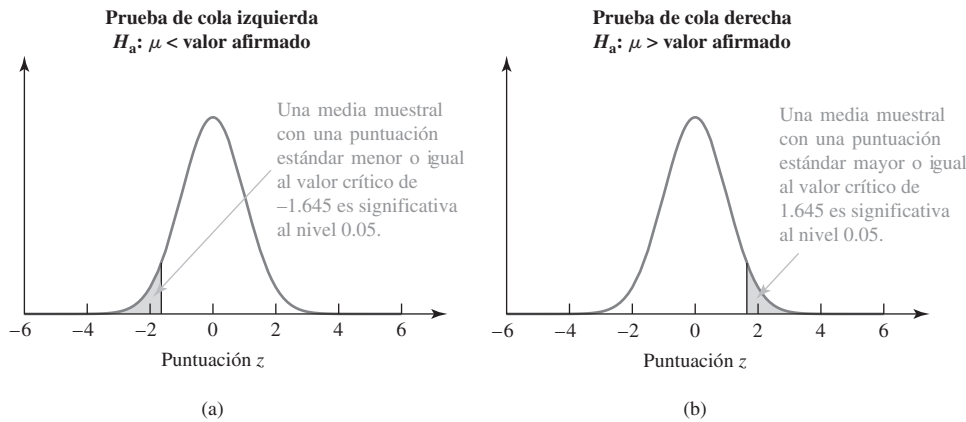


Figura 9.3 Estas gráficas ilustran el significado de los valores críticos para una hipótesis de una cola cuando se prueba por significancia en el nivel 0.05. La ubicación de los valores críticos corresponde al quinto percentil para pruebas de cola izquierda y el percentil 95o. para pruebas de cola derecha.

Decisiones con base en significancia estadística para pruebas de hipótesis de una cola

Decidimos si rechazamos o no la hipótesis nula comparando la puntuación estándar (z) para una media muestral con los **valores críticos** para significancia a un nivel dado. La tabla 9.1 resume las decisiones para pruebas de hipótesis de una cola a los niveles de significancia de 0.05 y de 0.01.

Tabla 9.1 Prueba para significancia a los niveles 0.05 y 0.01

Tipo de prueba	Forma de H_a	Para nivel 0.05: rechazar H_0 si la puntuación estándar es	Para nivel 0.01: rechazar H_0 si la puntuación estándar es
Prueba de cola izquierda	$H_a: \mu < \text{valor afirmado}$	$z \leq -1.645$	$z \leq -2.33$
Prueba de cola derecha	$H_a: \mu > \text{valor afirmado}$	$z \geq 1.645$	$z \geq 2.33$

UN MOMENTO DE REFLEXIÓN

Suponga que, en pruebas de cola derecha, un estudio encuentra una media muestral con $z = 3$ y otro estudio encuentra una media muestral con $z = 10$. Ambas son significativas al nivel 0.01, pero ¿cuál resultado proporciona evidencia más fuerte para rechazar la hipótesis nula? Explique.

EJEMPLO 1 Prueba de significancia para Columbia College

Suponiendo que la hipótesis nula es verdadera y el salario medio inicial para los graduados de Columbia College es \$39 000, ¿es estadísticamente significativo encontrar una muestra en la que la media es sólo \$37 000? Con base en su respuesta, ¿debemos rechazar o no la hipótesis nula?

Solución Recuerde que la prueba de hipótesis es de cola izquierda (ya que la hipótesis alternativa tiene la forma “menor que” de $\mu < \$39\,000$) y ya encontramos que una media muestral $\bar{x} = \$37\,000$ tiene una puntuación estándar de -3.25 . Este resultado es significativo al nivel 0.05 porque la puntuación estándar es menor que el valor crítico de $z = -1.645$. De hecho, es muy significativo que en el nivel 0.01 la puntuación estándar es menor que -2.33 . Por tanto, tenemos razón para rechazar la hipótesis nula y concluir que los directivos de Columbia College exageraron el salario medio inicial de sus graduados.

Determinación del valor P

Como estudiamos en la sección 9.1, podemos utilizar los valores P para ser más precisos acerca de la significancia de un resultado. Recuerde que el valor P es la probabilidad de encontrar una media muestral tan extrema como la encontrada, bajo el supuesto que la hipótesis nula es verdadera. Para las pruebas de hipótesis que analizamos en este capítulo, por lo general encontramos el valor P de la puntuación estándar de la media muestral. La figura 9.4 muestra la idea. El área total bajo la curva para la distribución de medias muestrales se define como 1, por lo que podemos interpretar las áreas bajo la curva como probabilidades. Para una prueba de cola izquierda (figura 9.4a), la probabilidad de encontrar una media muestral menor o igual a algún valor particular es simplemente el área debajo de la curva a la izquierda de la media muestral. Para una prueba de cola derecha (figura 9.4b), la probabilidad es el área debajo de la curva a la derecha de la media muestral.

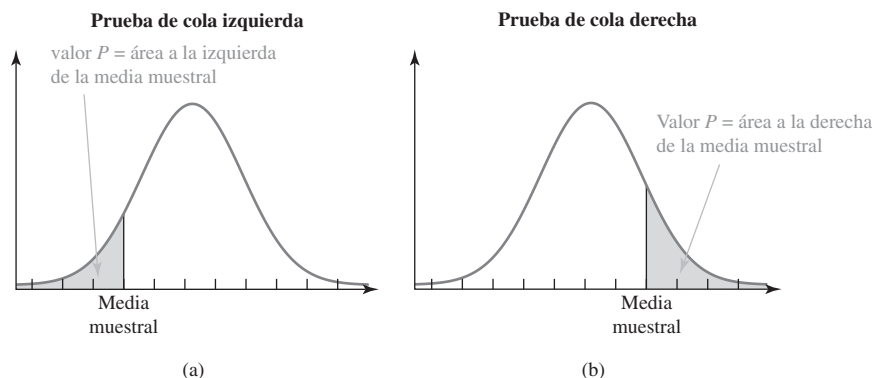


Figura 9.4 El valor P para pruebas de hipótesis de una cola corresponde al área debajo de la curva de distribución de muestreo. Una vez que calculamos la puntuación estándar para una media muestral, podemos encontrar el área correspondiente en las tablas de puntuación estándar como las del apéndice A. (Nota: el apéndice A lista las áreas a la izquierda de cada puntuación estándar; por tanto, para determinar el área a la derecha de la media muestral en pruebas de cola derecha, se requiere restar de 1, el valor dado en la tabla).

El valor P . . . con frecuencia es usado como un indicador del grado de rechazo en que la hipótesis nula merece tenerse en cuenta.

—Robert Abelson, *Statistical as a Principled Argument*

Como un ejemplo, determinemos el valor P para la prueba de hipótesis de Columbia College. Ya hemos determinado que la media muestral $\bar{x} = \$37\,000$ tiene una puntuación estándar de $z = -3.25$. Si revisa el apéndice A verá que esta puntuación estándar corresponde a un

“área acumulada a la izquierda” de 0.0006, por lo que éste es el valor P que estamos buscando. Este valor P pequeño proporciona fuerte evidencia contra la hipótesis nula. Asegúrese de notar que este valor P es menor que 0.01, por lo cual el resultado es significativo al nivel 0.01; el hecho que sea mucho menor que 0.01 nos dice que tenemos muy fuerte evidencia para rechazar la hipótesis nula.

Resumen de pruebas de una cola para medias poblacionales

Iniciamos esta sección con el objetivo de aprender a realizar los pasos 3 y 4 del proceso de cuatro pasos para pruebas de hipótesis en la página 376. Ahora hemos tenido éxito en hacer esto para pruebas de hipótesis de una cola para medias poblacionales. En resumen:

- Puesto que estamos tratando con medias poblacionales, la hipótesis nula tiene la forma $\mu = \text{valor afirmado}$. Para decidir si se rechaza o no se rechaza la hipótesis nula, debemos determinar si una muestra tan extrema como la encontrada en la prueba de hipótesis es probable o poco probable de ocurrir si la hipótesis nula es verdadera.
- Determinamos esta probabilidad a partir de la puntuación estándar, que calculamos con base en la fórmula

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

donde n es el tamaño de la muestra, $\bar{x}$ es la media muestral, μ es la media población afirmada por la hipótesis nula y σ es la desviación estándar poblacional. (En esta sección, por lo general aproximamos σ con la desviación estándar muestral, s ; una mejor aproximación se describe en la sección 10.1).

- Entonces, podemos evaluar la puntuación estándar de dos formas:
 1. Podemos evaluar su nivel de significancia estadística comparándola con los valores críticos dados en la tabla 9.1 (página 383).
 2. Podemos determinar su valor P con tablas de puntuaciones estándar como las del apéndice A. Para una prueba de cola izquierda, el valor P es el área bajo la curva normal a la izquierda de la puntuación estándar; para una prueba de cola derecha, es el área bajo la curva normal a la derecha de la puntuación estándar.
- Si el resultado es estadísticamente significativo al nivel elegido (por lo común nivel de significancia de 0.05 o 0.01) rechazamos la hipótesis nula. Si no es estadísticamente significativo, no rechazamos la hipótesis nula.

EJEMPLO 2 Millaje medio de automóviles rentados (revisado)

Recuerde el caso del propietario de la flota de automóviles en renta (ejemplo 5 de la sección 9.1) quien sospecha que el millaje medio anual de sus automóviles es mayor que la media nacional de 12 000 millas. Él selecciona una muestra aleatoria de $n = 225$ automóviles y calcula la media muestral $\bar{x} = 12 375$ millas y la desviación estándar muestral $s = 2 415$ millas. Determine el nivel de significancia estadística y el valor P para esta prueba de hipótesis, e interprete sus hallazgos.

Solución Para determinar la significancia y el valor P , primero debemos calcular la puntuación estándar para la media muestral de $\bar{x} = 12 375$ millas. Se nos dio la media poblacional afirmada ($\mu = 12 000$) millas y el tamaño de la muestra ($n = 225$); no conocemos la desviación estándar poblacional, σ , pero supondremos que es igual a la desviación estándar muestral y hacemos $\sigma = 2 415$ millas. Encontramos

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{12 375 - 12 000}{2 415 / \sqrt{225}} = 2.33$$

Esta puntuación estándar es mayor que el valor crítico de $z = 1.645$ al nivel 0.05 y es igual al valor de $z = 2.33$ para la significancia al nivel 0.01, dándonos una razón sólida para rechazar la hipótesis nula y concluir que en realidad el promedio de la flota de automóviles rentados es

mayor que el promedio nacional. Podemos encontrar el valor P del apéndice A, el cual muestra que el área a la *izquierda* de una puntuación estándar de $z = 2.33$ es 0.9901; puesto que es una prueba de cola derecha, la probabilidad es el área a la *derecha* (vea la figura 9.4b), que es $1 - 0.9901 = 0.0099$. Por tanto, el valor P es 0.0099, que es muy cercano a 0.01.

Pruebas de dos colas

Las mismas ideas básicas se aplican a las pruebas de hipótesis de dos colas en que la hipótesis alternativa tiene la forma “no es igual” de $H_a: \mu \neq$ valor afirmado. Sin embargo, los valores críticos y los cálculos para los valores P son un poco diferentes para pruebas de dos colas.

Como siempre, la prueba de hipótesis es significativa al nivel 0.05 si la probabilidad de encontrar un resultado tan extremo como el encontrado en realidad es 0.05 o menos. Para pruebas de una cola, una probabilidad de 0.05 corresponde a puntuaciones estándar en el quinto percentil para pruebas de cola izquierda, y al 95o. percentil para pruebas de cola derecha (vea la figura 9.3). Sin embargo, en las pruebas de dos colas un valor “tan extremo como el realmente encontrado” puede estar a *cualquier* lado, izquierdo o derecho, de la distribución muestral (figura 9.5). Por tanto, una probabilidad de 0.05, o 5%, corresponde a puntuaciones estándar, ya sea en el primero 2.5% de la distribución muestral a la izquierda o en el último 2.5% de la derecha. Del apéndice A, el percentil 2.5o. corresponde a una puntuación estándar de -1.96 y el percentil 97.5o. corresponde a una puntuación estándar de 1.96 . Estas puntuaciones estándar se vuelven valores críticos para pruebas de dos colas que sean significativas al nivel 0.05. De manera similar, los valores críticos para dos colas con significancia al nivel 0.01 corresponden a puntuaciones estándar del 0.5o. y 99.5o. percentiles; del apéndice A, estos valores son -2.575 y 2.575 , respectivamente.

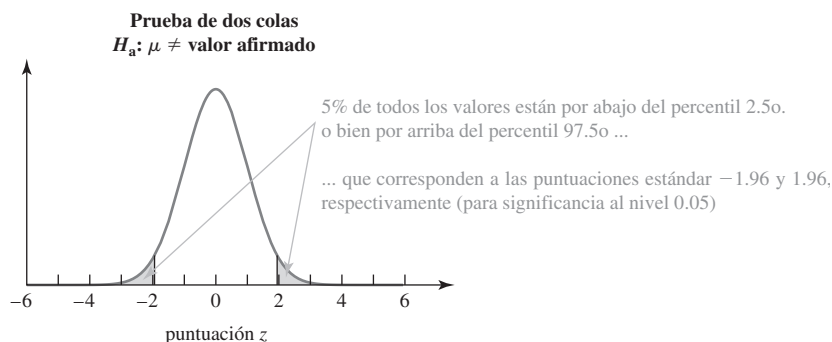


Figura 9.5 Para pruebas de dos colas, los valores críticos para la significancia al nivel 0.05 corresponden a los percentiles 2.5o. y 97.5o. (en comparación con el 5o. y 95o. percentiles para pruebas de una cola).

Consideraciones similares se aplican a valores P para pruebas de dos colas. Recuerde que para una prueba de una cola, el valor P es el área debajo de la curva a la izquierda de la media muestral para una prueba de cola izquierda o a la derecha de la media muestral para una prueba de cola derecha. Puesto que una prueba de dos colas nos pide considerar extremos en ambos lados de la media afirmada, el valor P para una prueba de dos colas debe ser el doble de lo que sería si la prueba fuese de una cola.

Prueba de dos colas ($H_a: \mu \neq$ valor afirmado)

Significancia estadística: una prueba de dos colas es significativa al nivel 0.05 si la puntuación estándar de la media muestral está en o debajo de un valor crítico de -1.96 o en o arriba de un valor crítico de 1.96 . Para una significancia en el nivel 0.01, los valores críticos son -2.575 y 2.575 .

Valores P : para encontrar el valor P para una media muestral en una prueba de dos colas, primero utilice la puntuación estándar de la media muestral para determinar el valor P suponiendo que la prueba es de una cola y luego duplique este valor para determinar el valor P para la prueba de dos colas.

Un ejemplo ayudará a clarificar estas ideas. Considere una compañía que busca asegurarse que sus tabletas de aspirinas de “500 miligramos” en realidad contienen 500 miligramos de aspirina. Si las tabletas contienen menos de 500 miligramos, los consumidores no obtendrán la dosis que se anunció. Si las tabletas contienen más de 500 miligramos, los consumidores están obteniendo demasiada medicina. La hipótesis nula dice que la media poblacional del contenido de aspirina es 500 miligramos:

$$H_0: \mu = 500 \text{ miligramos}$$

La compañía está interesada en la posibilidad de que el peso medio sea *o bien* menor *o* mayor que 500 miligramos. Puesto que la compañía está interesada en posibilidades de ambos lados de la media afirmada de 500 miligramos, tenemos una prueba de dos colas en que la hipótesis alternativa es

$$H_a: \mu \neq 500 \text{ miligramos}$$

Suponga que la compañía selecciona una muestra aleatoria de $n = 100$ tabletas y encuentra que tienen un peso medio de $\bar{x} = 501.5$ miligramos; además, suponga que la desviación estándar poblacional es $\sigma = 7.0$ miligramos. Entonces la puntuación estándar (z) para esta media muestral es

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{501.5 - 500}{7/\sqrt{100}} = 2.14$$

Esta puntuación estándar está por arriba del valor crítico de 1.96 para una prueba de dos colas, por lo que es significativa al nivel 0.05. (No es significativa al nivel 0.01, ya que no está por arriba del valor crítico de 2.575). Esto nos da buena razón para rechazar la hipótesis nula y concluir que el peso medio de tabletas de “500 miligramos” es diferente de 500 miligramos.

Encontramos el valor P usando el apéndice A, el cual muestra que el área a la *izquierda* de una puntuación de $z = 2.14$ es 0.9838. Por tanto, si ésta fuese una prueba de una cola, el valor P sería el área a la *derecha*, o $1 - 0.9838 = 0.0162$. Pero puesto que es una prueba de dos colas, duplicamos este número para encontrar que el valor P es $2 \times 0.0162 = 0.0324$. En otras palabras, *si* la hipótesis nula es verdadera, la probabilidad de sacar una muestra tan extrema como la encontrada es 0.0324.

EJEMPLO 3 ¿Cuál es la temperatura media del cuerpo humano?

Considere otra vez el estudio en el que los investigadores de la Universidad de Maryland midieron la temperatura del cuerpo en una muestra de $n = 106$ adultos saludables (vea el ejemplo 5 en la sección 8.2), determinando una temperatura media corporal de $\bar{x} = 98.20^\circ\text{F}$ con una desviación estándar muestral de $s = 0.62^\circ\text{F}$. Supondremos que la desviación estándar poblacional es la misma que la desviación estándar muestral. Determine si esta muestra proporciona evidencia para rechazar la creencia común de que la temperatura media del cuerpo humano es $\mu = 98.6^\circ\text{F}$.

Solución La hipótesis nula es la afirmación de que la temperatura media del cuerpo humano es 98.6°F , o $H_0: \mu = 98.6^\circ\text{F}$. Nos piden probar la hipótesis alternativa de que la temperatura media del cuerpo humano *no* es 98.6°F , o $H_a: \mu \neq 98.6^\circ\text{F}$. La forma “no es igual a” de la hipótesis alternativa nos dice que necesitamos una prueba de dos colas. Nos dan el tamaño de la muestra $n = 106$, la media muestral $\bar{x} = 98.20^\circ\text{F}$, y la desviación estándar poblacional supuesta $\sigma = 0.62^\circ\text{F}$. La puntuación estándar de la media muestral es

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{98.20 - 98.60}{0.62/\sqrt{106}} = -6.64$$

Esta puntuación estándar es mucho menor que los valores críticos de -1.96 para significancia en el nivel 0.05 y -2.575 para significancia en el nivel 0.01 (figura 9.6). Puesto que el resultado es significativo al nivel 0.01, proporciona evidencia contra la hipótesis nula. Por tanto, rechazamos la hipótesis nula y concluimos que la temperatura media del cuerpo humano *no* es igual a 98.6°F . (En este caso, la puntuación estándar de -6.64 es tan extrema que no podemos encontrar el valor P preciso del apéndice A; un cálculo muestra que el valor P es alrededor de 3×10^{-11}).

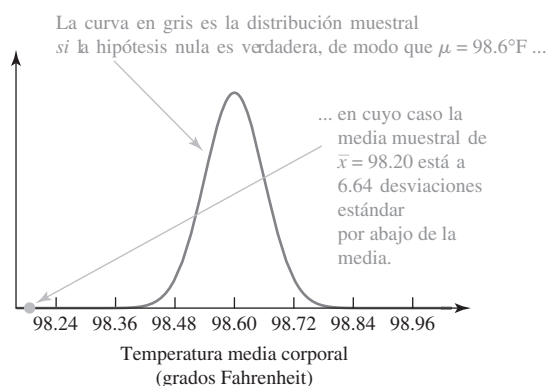


Figura 9.6 La hipótesis nula establece que la temperatura media corporal humana es 98.6°F . Si esto es verdadero, entonces la distribución de medias muestrales tiene una media de 98.6°F . Con base en el tamaño de la muestra y la desviación estándar, la media muestral de 98.20°F está a más de 6 desviaciones estándar por debajo de la temperatura media afirmada, lo cual hace probable parecer que la afirmación es errónea y la temperatura media corporal humana *no* es igual a 98.6°F .

Errores comunes en pruebas de hipótesis

Ahora hemos cubierto todas las pruebas de hipótesis fundamentales, así como los métodos específicos usados en pruebas de hipótesis para medias poblacionales. Sin embargo, aunque si una prueba de hipótesis es llevada a cabo de manera correcta, dos tipos de errores pueden afectar la conclusión. Para entender estos errores, recuerde la analogía legal al final de la sección 9.1, en que la hipótesis nula es H_0 : el acusado es inocente. Un tipo de error ocurre si concluimos que el acusado es culpable cuando, en realidad, es inocente. En este caso, la hipótesis nula (inocente) ha sido rechazada de manera incorrecta. El otro tipo de error ocurre si encontramos que el acusado *no es culpable* cuando en realidad es culpable. En este caso, hemos fallado erróneamente al rechazar la hipótesis nula.

Para un ejemplo estadístico, considere otra vez la prueba de la compañía farmacéutica que asegura que la cantidad media de aspirina en sus tabletas es 500 miligramos ($H_0: \mu = 500$ miligramos). Al sacar una conclusión de la prueba, la compañía podría cometer los dos tipos de errores siguientes:

- La compañía podría rechazar la hipótesis nula y concluir que la cantidad media no es 500 miligramos cuando sí lo es. Este error puede llevar a desperdiciar tiempo y dinero tratando de arreglar un proceso que no está mal. Un error de este tipo, en el que H_0 es *rechazada erróneamente*, se denomina **error de tipo I**.
- La compañía podría equivocarse en rechazar la hipótesis nula cuando, en realidad, la cantidad media de aspirina no es 500 miligramos. En este caso, la compañía distribuiría tabletas que tienen demasiada o muy poca aspirina; los consumidores podrían sufrir o demandar a la compañía. Un error de este tipo, en el que *de manera errónea se fracasa en rechazar H_0* , se denomina **error de tipo II**.

La tabla 9.2 resume los cuatro casos posibles.

Tabla 9.2 Tabla de decisión para H_0 y H_a			
		Realidad	
		H_0 verdadera	H_a verdadera
Decisión	Rechazar H_0	Error de tipo I	Decisión correcta
	No rechazar H_0	Decisión correcta	Error de tipo II

Si piensa acerca de esto se dará cuenta que hay una relación importante entre el nivel de significancia de una prueba de hipótesis y los errores de tipo I (rechazar erróneamente H_0): *el*

nivel de significancia es la probabilidad de cometer el error de tipo I. Por ejemplo, si realizamos una prueba de hipótesis usando un nivel de significancia de 0.05, existe una probabilidad de 0.05 de cometer el error de rechazar la hipótesis nula cuando en realidad es verdadera. Si realizamos la prueba a un nivel de significancia de 0.01, existe una probabilidad de 0.01 de rechazar de manera errónea la hipótesis nula. (La probabilidad de un error de tipo II también puede ser cuantificada pero está fuera del alcance de este libro).

EJEMPLO 4 Errores en la prueba de la temperatura corporal

Considere la hipótesis nula del ejemplo 3: que la temperatura media del cuerpo es igual a 98.6°F ($H_0: \mu = 98.6^\circ\text{F}$).

- ¿Qué decisiones correctas son posibles con esta hipótesis nula?
- En este caso, explique el significado de los errores de tipo I y de tipo II.
- En el ejemplo 3 rechazamos la hipótesis nula. ¿Cuál es la probabilidad que cometiésemos un error al hacerlo?

Solución

- Cualquier hipótesis tiene dos posibles decisiones correctas. En este caso, una decisión correcta ocurre si la temperatura media correcta del cuerpo en realidad es 98.6°F y *no* rechazamos H_0 . La otra decisión correcta ocurre si la temperatura media del cuerpo *no* es 98.6°F y rechazamos H_0 .
- Un error de tipo I ocurre si rechazamos H_0 cuando en realidad es verdadera. En este caso ocurre un error de tipo I si la temperatura media del cuerpo en realidad es 98.6°F pero concluimos que no lo es. Un error de tipo II ocurre si no rechazamos H_0 cuando en realidad es falsa. En este caso, un error de tipo II ocurre si la temperatura media del cuerpo no es 98.6°F pero nos equivocamos en llegar a esta conclusión.
- Rechazamos la hipótesis nula a un nivel de significancia de 0.01. Por tanto, existe una probabilidad de 0.01 de que cometamos un error de tipo I de rechazar una hipótesis nula que en realidad es verdadera.

Sesgo al seleccionar las hipótesis

Como hemos visto, una prueba de hipótesis bien llevada a cabo sigue un proceso muy estricto que lleva a minimizar la posibilidad de contaminar la prueba con sesgo. Sin embargo, el sesgo puede introducirse de muchas maneras, y en particular puede ocurrir en la selección inicial de las hipótesis. Considere una situación en la que una fábrica está investigando la liberación de contaminantes en un río cercano a un nivel que pueda (o no) exceder el máximo permitido por el gobierno. La elección lógica para la hipótesis nula es que el nivel medio real de contaminante es igual al nivel máximo permitido:

$$H_0: \text{nivel medio de contaminantes} = \text{nivel máximo permitido}$$

Sin embargo, esto nos deja con dos elecciones razonables para la hipótesis alterna, cada una de las cuales introduce un sesgo en la conclusión final.

Si los *administradores de la fábrica* realizan la prueba, estarían inclinados a asegurar que el nivel medio de contaminación es *menor que* el nivel máximo permitido. Esto es, elegirían

$$H_a: \text{nivel medio de contaminación} < \text{nivel máximo permitido}$$

Con esta hipótesis alternativa, los dos posibles resultados son

- Rechazar H_0 , en cuyo caso concluimos que el nivel medio de contaminantes es menor que el máximo permitido. Este resultado le gustaría a los administradores de la fábrica.
- No rechazar H_0 , en cuyo caso la prueba no es concluyente.

En otras palabras, eligiendo la prueba de hipótesis de cola izquierda (“menor que”), los administradores de la fábrica aseguran que en el peor de los casos terminaremos sin evidencia de que ellos excedieron el máximo permitido, aunque es posible que la conclusión esté a su favor.

Ahora supongamos que un *grupo ambiental* realiza la prueba. Puesto que el grupo sospecha que la fábrica está violando los estándares del gobierno, su elección para la hipótesis alternativa que el nivel de contaminación es *mayor que* el máximo permitido, o

$$H_a: \text{nivel medio de contaminantes} > \text{nivel máximo permitido}$$

Con esta elección de hipótesis alternativa, el rechazo de H_0 implica que la fábrica ha violado los estándares, mientras que no rechazar H_0 nuevamente no es concluyente. En otras palabras, la elección de una prueba de cola derecha (“mayor que”) crea una situación en que la fábrica puede ser encontrada en violación, pero no puede probarse que esté cumpliendo.



EJEMPLO 5 Mina de oro

El éxito de las minas de metales preciosos depende de la pureza (o grado) del mineral eliminado y el precio del mercado para el metal. Suponga que la pureza del mineral de oro debe ser al menos de 0.5 onzas de oro por tonelada de mineral, para mantener abierta una mina. Muestras de mineral de oro son usadas para estimar la pureza del mineral para toda la mina. Analice el impacto de los errores de tipo I y II en dos de las posibles hipótesis alternativas:

$$H_a: \text{pureza} < 0.5 \text{ onzas por tonelada}$$

$$H_a: \text{pureza} > 0.5 \text{ onzas por tonelada}$$

Solución Para el caso de cola izquierda, la hipótesis nula y la alternativa son

$$H_0: \text{pureza} = 0.5 \text{ onzas por tonelada}$$

$$H_a: \text{pureza} < 0.5 \text{ onzas por tonelada}$$

Las dos decisiones posibles en este caso son:

- Rechazar H_0 , que significa concluir que la pureza del mineral es menor que el necesario para mantener abierta la mina, por lo que la mina es cerrada.
- No rechazar H_0 , lo que significa que tenemos evidencia insuficiente para concluir que la pureza de la mina es menor que el necesario, por lo que la mina permanece abierta.

Un error de tipo I (rechazar de manera errónea una H_0 verdadera) significa que la mina es cerrada cuando, en realidad, la pureza del mineral de oro es suficiente para operar la mina. Para los operadores de la mina esto significa la pérdida de utilidades potenciales de la mina; para los empleados significa pérdida innecesaria de trabajo. Un error de tipo II (equivocarse en rechazar una H_0 falsa) significa que la mina continúa en operación, pero en realidad no es rentable.

Ahora suponga que la mina se mantiene abierta sólo si la pureza es mayor a 0.5 onzas por tonelada. Para este caso de cola derecha, las hipótesis nula y alternativa son

$$H_0: \text{pureza} = 0.5 \text{ onzas por tonelada}$$

$$H_a: \text{pureza} > 0.5 \text{ onzas por tonelada}$$

Las dos decisiones posibles son

- Rechazar H_0 , que significa concluir que la pureza del mineral es *más* que el necesario para mantener abierta la mina, por lo que la mina permanece abierta.
- No rechazar H_0 , lo que significa que tenemos evidencia insuficiente para concluir que la pureza de la mina es lo suficientemente alta para mantener abierta la mina, por lo que la mina es cerrada.

Un error de tipo I (rechazar de manera incorrecta una H_0 verdadera) significa dejar abierta la mina cuando en realidad no es rentable. Un error de tipo II (no rechazar una H_0 falsa) significa cerrar la mina cuando en realidad es rentable, dejando a los empleados sin trabajo cuando no hay razón suficiente.

Sección 9.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Notación.** Cuando realizamos pruebas de hipótesis para una media poblacional, es muy importante entender claramente lo que representan n , $\bar{x}$, s , σ y μ . Describa qué indica cada una de estas variables.
 - Pulso.** A una estudiante graduada se le ha asignado probar la afirmación de que el pulso de los instructores de la universidad es mayor que 70 latidos por minuto. Ella mide el pulso de 225 instructores en su universidad, y encuentra que hay suficiente evidencia para apoyar la afirmación. ¿Qué es incorrecto en su proyecto?
 - Tipos de errores.** ¿Qué es el error de tipo I? ¿Qué es el error de tipo II?
 - Significancia estadística.** En una prueba de la dieta Atkins se obtuvo un valor P de 0.004 cuando se comprobó la afirmación de que había una pérdida con la dieta. El peso medio perdido después de 12 meses fue 2.1 libras. ¿El peso medio perdido de 2.1 libras es estadísticamente significativo? ¿El peso medio perdido de 2.1 libras tiene significancia práctica? ¿Por qué sí o por qué no?
- ¿Tiene sentido?** Para los ejercicios 5 al 12 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.
- Prueba de hipótesis.** En una prueba de hipótesis el valor P es el mismo que el nivel de significancia.
 - Valor P .** Al probar una afirmación respecto a una media poblacional, un estadístico de prueba de $z = 0.20$ corresponde a un valor P más grande que un estadístico de prueba de $z = 2.00$.
 - Valor P .** En una prueba de hipótesis, un valor P de 0.999 indica que debe respaldar la hipótesis alternativa.
 - Hipótesis alternativa.** Al probar la afirmación de que la puntuación media del CI de reclusos en prisión es menor que 100, la hipótesis alternativa se expresa como $\bar{x} < 100$.
 - Significancia.** El nivel de significancia en una prueba de hipótesis es la probabilidad de cometer un error de tipo I.
 - Error de tipo I.** Cuando un grupo de consumidores está probando la afirmación de que la cantidad media de aspirina en tabletas es 350 miligramos, es extremadamente importante no rechazar erróneamente una hipótesis nula verdadera. Así, es mejor elegir 0.01 que 0.05 para el nivel de significancia.
 - Nivel de significancia.** Puesto que el nivel de significancia es la probabilidad de cometer un error de tipo I, es prudente

seleccionar un nivel de significancia de cero, de modo que no haya probabilidad de cometer ese error.

- Mnemónico.** Una regla mnemotécnica para interpretar el valor P en una prueba de hipótesis es: “Si la P (valor) es baja, la nula no trabaja”.

Conceptos y aplicaciones

¿Se respaldan las hipótesis alternativas? En los ejercicios 13 al 20 encuentre el valor de la puntuación estándar, z , y determine si la hipótesis alternativa es respaldada a un nivel de significancia de 0.05.

- $H_a: \mu > 100$, $n = 100$, $\bar{x} = 102$, $\sigma = 15$
- $H_a: \mu > 69.0$, $n = 400$, $\bar{x} = 75.2$, $\sigma = 12.0$
- $H_a: \mu < 98.6$, $n = 64$, $\bar{x} = 98.2$, $\sigma = 0.6$
- $H_a: \mu < 1\,518$, $n = 81$, $\bar{x} = 1\,516$, $\sigma = 14$
- $H_a: \mu \neq 37.40$, $n = 400$, $\bar{x} = 37.30$, $\sigma = 0.80$
- $H_a: \mu \neq 24.7$, $n = 100$, $\bar{x} = 24.0$, $\sigma = 1.9$
- $H_a: \mu \neq 1\,050$, $n = 49$, $\bar{x} = 1\,058$, $\sigma = 20$
- $H_a: \mu \neq 777$, $n = 900$, $\bar{x} = 779$, $\sigma = 42$

Determinación de valores P . En los ejercicios 21 al 34 utilice la tabla 5.1 para encontrar el valor P que corresponde a la puntuación estándar z , y determine si la hipótesis alternativa es respaldada a un nivel 0.05 de significancia.

- $z = -0.4$ para $H_a: \mu < 25$
- $z = -2.4$ para $H_a: \mu < 727$
- $z = 1.9$ para $H_a: \mu > 36.35$
- $z = 1.5$ para $H_a: \mu > 0.227$
- $z = -1.6$ para $H_a: \mu \neq 0.389$
- $z = -2.0$ para $H_a: \mu \neq 172$
- $z = 1.7$ para $H_a: \mu \neq 75$
- $z = 1.9$ para $H_a: \mu \neq 25.7$
- $z = 2.7$ para $H_a: \mu > 19.4$
- $z = 3.5$ para $H_a: \mu > 75$
- $z = -2.1$ para $H_a: \mu < 1\,007$
- $z = -1.1$ para $H_a: \mu < 149.6$
- $z = 0.15$ para $H_a: \mu \neq 90.3$
- $z = 3.5$ para $H_a: \mu \neq 1\,022$

35. Interpretación de valores P . Suponga que está probando una hipótesis alternativa de la forma $H_a: \mu >$ valor afirmado. Si la media muestral tiene una puntuación estándar de $z = -1.0$, ¿qué concluye? ¿Por qué no es necesario realizar la prueba de hipótesis formal?

36. Interpretación de valores P . Suponga que está probando una hipótesis alternativa de la forma $H_a: \mu >$ valor afirmado. Si la media muestral tiene una puntuación estándar de $z = 0.5$, ¿qué concluye? ¿Por qué no es necesario realizar la prueba de hipótesis formal?

Pruebas de hipótesis para medias. Para los ejercicios 37 al 50 utilice un nivel de significancia 0.05 y realice una prueba de hipótesis usando el proceso de cuatro pasos descrito en el texto. Asegúrese de establecer su conclusión.

37. Encuesta Roper. Una encuesta Roper utilizó una muestra de 100 propietarios de automóvil seleccionados de manera aleatoria. En la muestra, el tiempo medio de posesión de un solo automóvil fue de 7.01 años, con una desviación estándar de 3.74 años. Pruebe la afirmación hecha por el propietario de una gran concesionaria de que el tiempo medio de posesión para todos los automóviles es menor que 7.5 años.

38. Tiempo en el hospital. De acuerdo con un estudio realizado por los Centros para el Control de Enfermedades, la media nacional de estancia en hospital después del alumbramiento es 2.0 días. Al revisar los registros en su hospital, una administradora calcula que la media de estancia en el hospital para una muestra de 81 mujeres que dieron a luz es 2.2 días con una desviación estándar de 1.2 días. Suponiendo que las pacientes representan una muestra aleatoria de la población, pruebe la afirmación de que este hospital mantiene más tiempo a las nuevas mamás que el promedio nacional. ¿Cómo cambia el resultado si la desviación estándar de la muestra es 0.8 días?

39. Empaquetamiento. Los fabricantes de una marca líder de pasta dental necesitan asegurarse de que la cantidad de pasta dental en cada paquete de “16 onzas” no se sobrepase demasiado (lo cual tendería a desperdiciar producto) ni que le falte (lo cual llevaría a la insatisfacción de los clientes). Los paquetes de pasta dental en una muestra de $n = 144$ tienen un peso medio de 15.8 onzas con una desviación estándar de 1.6 onzas. Pruebe la afirmación de que los paquetes de “16 onzas” en realidad tienen un peso medio de 16 onzas.

40. Fabricación de motocicletas. Los fabricantes de motocicletas deben producir ejes delanteros que cumplan con dimensiones especificadas. En particular, los diámetros de los ejes deben ser de 8.50 centímetros. Los ejes en una muestra de $n = 64$ ejes tienen un diámetro medio de 8.56 centímetros con una desviación estándar de 0.24 centímetros. Pruebe la afirmación de que los ejes delanteros tienen el diámetro medio especificado.

41. Cantidades de medicina. La medicina en crema Dozenol lista 600 miligramos de acetaminofén por onza como un ingrediente activo. La Administración de Alimentos y Dro-

gas (AAD) prueba 65 muestras de una onza de la medicina y encuentra que la cantidad media de acetaminofén para la muestra es 589 miligramos con una desviación estándar de 21 miligramos. Pruebe la afirmación de la AAD de que la medicina no contiene la cantidad requerida de acetaminofén.

42. Ingreso. Recién el ingreso medio en hogares en el condado de Wasatch, Utah, fue \$41 045 con una desviación estándar de \$1 605. Suponga que los 13 267 residentes del condado representan una muestra aleatoria de todos los residentes de Utah. Con base en esta muestra, pruebe la afirmación de que el ingreso medio de los hogares en Utah difiere del ingreso medio nacional de \$35 100.

43. Consumo de combustible. De acuerdo con la Administración de Información de Energía (datos de la Administración de Autopistas Federales), el promedio de millaje de todos los automóviles es 21.4 millas por galón. Para una muestra aleatoria de 40 camionetas, el millaje medio es 19.8 millas por galón con una desviación estándar de 3.5 millas por galón. Pruebe la afirmación de que el millaje medio de todas las camionetas es menor que 21.4 millas por galón.

44. Bola de béisbol. Se obtuvo una muestra aleatoria de 40 bolas nuevas de béisbol. Cada una se dejó caer sobre una superficie de concreto y la altura de los rebotes tiene una media de 92.67 pulgadas y una desviación estándar de 1.79 pulgadas (con base en datos de *USA Today*). Pruebe la afirmación de que las bolas nuevas de béisbol tienen una altura media de rebote que es menor que la altura de rebote medio de 92.84 pulgadas encontrada para bolas viejas de béisbol.

45. Compradores compulsivos. Unos investigadores desarrollaron un cuestionario para identificar a compradores compulsivos. Una muestra aleatoria de 32 sujetos que se identifican como compradores compulsivos obtuvo una puntuación media en el cuestionario de 0.83 con una desviación estándar de 0.24 (con base en datos de “A Clinical Screener for Compulsive Buying”, de Faber y Guinn, *Journal of Consumer Research*, volumen 19). Pruebe la afirmación de que la población que se autoidentifica como comprador compulsivo tiene una media mayor que la media de 0.21 para la población general.

46. Peso de moneda. De acuerdo con el Departamento del Tesoro de Estados Unidos, el peso medio de una moneda de 25 centavos es 5.670 gramos. Una muestra aleatoria de 50 monedas tiene un peso medio de 5.622 gramos con una desviación estándar de 0.068 gramos. Pruebe la afirmación de que el peso medio de las monedas de 25 centavos en circulación es 5.670 gramos.

47. Peso al nacer. El peso medio al nacer de bebés hombres nacidos de 121 madres que tomaron un complemento vitamínico es 3.67 kilogramos con una desviación estándar de 0.66 kilogramos (con base en datos del Departamento de Salud del Estado de Nueva York). Pruebe la afirmación de que el peso medio al nacer de todos los bebés nacidos de madres que tomaron el complemento vitamínico es igual

a 3.39 kilogramos, que es la media para la población de todos los bebés hombres.

48. Millaje de un automóvil. De acuerdo con la Administración de Información de Energía (datos de la Administración de Autopistas Federales), el millaje promedio anual para automóviles en Estados Unidos es 11 725 millas. El propietario de una compañía de renta de automóviles selecciona una muestra aleatoria de 225 automóviles de su flota completa. Para la muestra, el millaje anual medio es 12 145 millas con una desviación estándar de 3 000 millas. Pruebe la afirmación de que el millaje medio de la flotilla es mayor que el promedio nacional.

49. Convicto por desfalco. Cuando 70 convictos por desfalco fueron seleccionados de manera aleatoria, se encontró que la duración media de las penas era 22.1 meses y la desviación estándar fue 8.6 meses (con base en datos del Departamento de Justicia de Estados Unidos). Jane Fleming está compitiendo por un puesto político en una plataforma que desea un tratamiento más severo a criminales convictos. Pruebe la afirmación de Jane de que las condenas en prisión para convictos por desfalco tienen una media menor que 24 meses.

50. Ingresos. De acuerdo con la Encuesta de Población Actual hecha por la Oficina del Censo, hombres con grado de licenciatura ganan un promedio de \$46 700. Suponga que una muestra aleatoria de 400 mujeres con grado de licenciatura tiene un ingreso medio de \$28 700 (la media nacional para mujeres) con una desviación estándar de \$16 500. Pruebe la afirmación de que la media de los ingresos de mujeres es menor que la media de los ingresos para hombres.

Errores de tipo I y de tipo II. En los ejercicios 51 a 54 se dan una hipótesis nula y una alternativa. Sin usar los términos “hipótesis nula” e “hipótesis alternativa”, identifique el error de tipo I y el error de tipo II.

51. H_0 : el paciente está libre de una enfermedad particular.

H_a : el paciente tiene la enfermedad.

52. H_0 : el acusado no es culpable.

H_a : el acusado es culpable.

53. H_0 : la lotería es justa.

H_a : la lotería está sesgada.

54. H_0 : la longitud media de un perno en el sistema de suspensión de los automóviles nuevos Audi es 3.456 centímetros.

H_a : la longitud media de un perno en el sistema de suspensión de los automóviles nuevos Audi no es igual a 3.456 centímetros.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 9 en www.aw.com/bbt.

55. Comparaciones con promedios nacionales. Seleccione diversas variables que sean, relativamente, sencillas de medir en una clase o muestra de estudiantes. Las variables deben incluir una cantidad que pueda ser promediada (por ejemplo, estatura, peso, tamaño de la familia, presión sanguínea, ritmo cardíaco, tiempo de reacción). Utilice internet u otras referencias para determinar los promedios nacionales para estas variables (por categorías de edad, si es adecuado). Recolecte información sobre las variables, usando una muestra aleatoria de por lo menos 50 individuos. Lleve a cabo las pruebas de hipótesis relevantes para determinar si es significativa la diferencia entre la media muestral y la media poblacional.

56. Datos de un condado. El reporte *Statistical Abstract of the United States* y la Encuesta de Población Actual proporcionan una fuente inagotable de estadísticas sociales, económicas y vitales a nivel de condado, estado y locales. Utilice sus sitios web para comparar los datos estatales con los datos nacionales de la manera siguiente:

- Seleccione una variable de interés que incluya una media para un estado particular (por ejemplo, la media del tamaño de los hogares en Illinois).
- Determine el valor nacional actual para esa variable (por ejemplo, la media del tamaño de los hogares).
- Seleccione un condado particular en el estado y obtenga los datos correspondientes para el condado, así como el tamaño de la muestra (por ejemplo, el tamaño medio de hogares en el condado de Cook, Illinois).
- Suponiendo que el condado es una muestra aleatoria del estado, pruebe la afirmación de que el estado está por arriba o por abajo del nivel nacional en términos de esa variable.
- Analice e interprete sus resultados. La prueba de hipótesis depende de que la muestra (el condado) sea una muestra aleatoria de la población (el estado). Asegúrese de analizar este factor en sus conclusiones.

57. Applet para pruebas de hipótesis. Usando un motor de búsqueda tal como Google, busque “prueba de hipótesis” y “applet”. Encuentre un applet y córralo. Describa cómo funciona el applet y qué ilustra.

58. Potencia de una prueba. Usando un motor de búsqueda, tal como Google, busque “potencia” de una prueba de hipótesis. Describa qué es la potencia de una prueba de hipótesis.

9.3 Pruebas de hipótesis para proporciones poblacionales

Ahora pasamos al problema de pruebas de hipótesis con *proporciones* (en lugar de medias). Todas las ideas de la sección anterior se aplican, excepto que necesitamos un método diferente para calcular la desviación estándar de la distribución muestral. Utilizamos un ejemplo para ilustrar el proceso.

Suponga que una candidata política encarga una encuesta antes de que se cierre una elección. Usando una muestra aleatoria de $n = 400$ probables votantes, la encuesta encuentra que 204 apoyan a la candidata. ¿La candidata debe estar confiada en que ganará? En el capítulo 8 analizamos cómo determinar el margen de error e intervalos de confianza para este tipo de encuesta. Ahora abordamos la pregunta como una prueba de hipótesis.

Para una prueba de hipótesis preguntamos si los resultados de la encuesta (que son las estadísticas muestrales) apoyan la hipótesis de que la candidata tiene más de 50% de los votos. Como es usual, representamos con p a la proporción de personas en la *población* votante que favorecen a la candidata, y denotamos con $\hat{p}$ a la proporción de personas en la *muestra* que están a favor de la candidata. Puesto que 204 de 400 personas en la muestra apoyan a la candidata, la proporción muestral es

$$\hat{p} = \frac{204}{400} = 0.510$$

Ahora podemos formular la hipótesis nula y la alternativa. Como es común, establecemos nuestra hipótesis nula como una igualdad:

$$H_0: p = 0.5 \text{ (50\% de los votantes a favor de la candidata)}$$

La candidata quiere saber si ella tiene respaldo de la mayoría, por lo que la hipótesis alternativa es de cola derecha:

$$H_a: p > 0.5 \text{ (más de 50\% de los votantes favorecen a la candidata)}$$

Cálculos para pruebas de hipótesis con proporciones

¿Cómo determinamos si existe evidencia suficiente en la muestra para rechazar la hipótesis nula? Siguiendo el proceso de pruebas de hipótesis de la página 376, vemos que hemos completado los primeros dos pasos (la formulación de las hipótesis y la recolección de datos muestrales). Ahora, al igual que como hicimos para las medias muestrales en la sección 9.2, debemos determinar la probabilidad de que la muestra resultante podría haber surgido por azar si la hipótesis nula es verdadera.

Procediendo como lo hicimos con medias muestrales, imaginamos la selección de muchas muestras de tamaño $n = 400$. Para cada muestra calculamos la proporción de personas que están a favor de la candidata. Otra vez, puesto que tenemos una muestra de tamaño razonablemente grande, la distribución de las proporciones muestrales debe ser muy cercana a una distribución normal. Bajo la suposición inicial de que la hipótesis nula es verdadera (que la proporción de personas en la población que están a favor de la candidata es 0.5), el pico de esta distribución será la proporción de la población afirmada por la hipótesis nula, $p = 0.5$. La desviación estándar de la distribución muestral está dada por la fórmula siguiente (la deducción está fuera del alcance de este libro):

$$\text{desviación estándar de la distribución de proporciones muestrales} = \sqrt{\frac{p(1-p)}{n}}$$

Para este caso, la desviación estándar es

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.5(1-0.5)}{400}} = 0.025$$

NOTA TÉCNICA

La fórmula para la desviación estándar de la distribución de proporciones muestrales es precisa sólo cuando $np \geq 5$ y $n(1-p) \geq 5$. Estas condiciones se cumplen para los problemas que consideramos en este libro.

La figura 9.7 muestra la distribución de las proporciones muestrales. Consideramos la proporción muestral de la encuesta ($\hat{p} = 0.510$) como un solo punto en esta distribución muestral. Muy parecido a las pruebas de hipótesis con medias (vea la figura 9.1), observamos lo siguiente:

- Si la muestra resultante está cerca del pico de la distribución muestral, entonces no tenemos razón para pensar que la hipótesis nula es incorrecta y no rechazamos la hipótesis nula.
- Si la muestra resultante está lejos del pico de la distribución muestral, entonces es más probable que la explicación sea que la distribución muestral no tiene un pico donde lo asegura la hipótesis nula, en cuyo caso rechazamos la hipótesis nula.

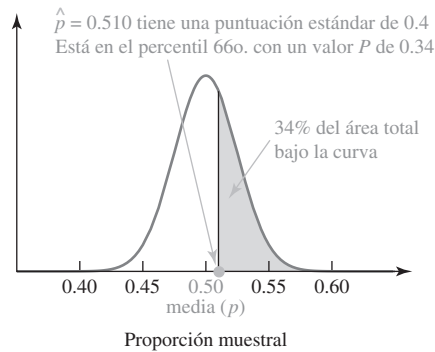


Figura 9.7 La distribución de proporciones muestrales para una encuesta de elecciones con una media de $p = 0.5$ y una desviación estándar de 0.025. La proporción muestral $\hat{p} = 0.510$ tiene una puntuación estándar de 0.4 y un valor P de 0.34.

De nuevo, como hicimos para medias muestrales, decidimos si la proporción muestral está “cerca” o “lejos” del pico de la distribución muestral (suponiendo que la hipótesis nula es verdadera) cuantificando la distancia con una puntuación estándar. La fórmula para la puntuación estándar (z) siempre tiene la misma forma general (vea la página 209); para la distribución de proporciones muestrales, la fórmula toma la forma

$$z = \frac{\text{proporción muestral} - \text{proporción poblacional}}{\text{desviación estándar de la distribución muestral}} = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

Ahora podemos calcular e interpretar la puntuación estándar para nuestro ejemplo actual. El tamaño de la muestra son las $n = 400$ personas entrevistadas, la proporción muestral es 51% de la muestra que apoya a la candidata ($\hat{p} = 0.510$) y la proporción poblacional es el $p = 0.5$ afirmado por la hipótesis nula. Por tanto, la puntuación estándar para la proporción muestral es

$$z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}} = \frac{0.510 - 0.5}{0.025} = 0.4$$

En otras palabras, la muestra usada en la encuesta tiene una proporción que está 0.4 desviaciones estándar por arriba del pico de la distribución de las proporciones muestrales. Como podemos ver en la figura 9.7, esta muestra resultante *no* es muy extrema, indicando que no debemos rechazar la hipótesis nula. Esto es, la candidata no puede estar confiada de obtener el apoyo de la mayoría.

Puntuación estándar para la proporción muestral en una prueba de hipótesis

Dado el tamaño de la muestra (n), la proporción muestral ($\hat{p}$) y la proporción poblacional afirmada (p), la puntuación estándar para la proporción muestral es

$$z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

A propósito...

En la elección presidencial de 1948 la mayoría de los encuestadores y periódicos predecían una victoria holgada para el republicano Dewey sobre el demócrata Truman. En realidad, Truman ganó por un margen de voto popular de 49.5% a 45.1%.

Niveles de significancia y valores P

Podemos usar un criterio más cuantitativo, usando un nivel de significancia o valor P . Iniciamos con la significancia. Puesto que el ejemplo de la encuesta de la elección utiliza una prueba de cola derecha, encontramos los valores críticos para significancia en la tabla 9.1 (página 383). La puntuación estándar de $z = 0.4$ *no* es mayor que el valor crítico de $z = 1.645$ para una significancia al nivel 0.05, confirmando que el resultado no debe hacernos rechazar la hipótesis nula. Para encontrar el valor P , usamos las tablas en el apéndice A. Las tablas muestran que el área a la *izquierda* de una puntuación estándar de $z = 0.4$ es 0.6554. Puesto que es una prueba de cola derecha, restamos este valor de 1 para determinar el área a la derecha (mostrada en la figura 9.7), que es el valor P para esta prueba de hipótesis. Esto es, el valor P es $1 - 0.6554 = 0.3446$, lo cual indica que si la hipótesis nula es verdadera, hay más de 0.34 de posibilidades de seleccionar aleatoriamente una muestra tan extrema como la encontrada en esta encuesta. Con tan alta probabilidad de sacar tal muestra por azar, no tenemos razón para rechazar la hipótesis nula, por lo que la candidata no puede suponer que ella tiene el apoyo de más de 50% de los votantes.

Resumen de pruebas de hipótesis con proporciones

Ahora podemos resumir el procedimiento para realizar una prueba de hipótesis con una proporción poblacional. Asegúrese de observar que aún estamos siguiendo el proceso general de cuatro pasos para pruebas de hipótesis de la página 376 y aquí nos centramos en lo específico de tratar con proporciones.

- Puesto que tratamos con proporciones poblacionales, la hipótesis nula tiene la forma $p =$ valor afirmado. Para decidir si se rechaza o no se rechaza la hipótesis nula, debemos determinar si una muestra tan extrema como la encontrada en la prueba de hipótesis es probable o poco probable que ocurra si la hipótesis nula es verdadera.
- Determinamos esta probabilidad de la puntuación estándar (z) de la proporción muestral, que calculamos de la fórmula

$$z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

donde n es el tamaño de la muestra, $\hat{p}$ es la proporción muestral y p es la proporción poblacional afirmada por la hipótesis nula.

- Entonces aplicamos la puntuación estándar de las dos formas que utilizamos para pruebas de hipótesis con medias:
 1. Podemos evaluar su nivel de significancia estadística comparando la puntuación estándar con los valores críticos dados en la tabla 9.1 (página 383) para pruebas de una cola y el recuadro de la página 386 para pruebas de dos colas.
 2. Podemos determinar su valor P con tablas de puntuaciones estándar como las del apéndice A. Para pruebas de cola izquierda, el valor P es el área debajo de la curva normal a la izquierda de la puntuación estándar; para una prueba de cola derecha, es el área debajo de la curva normal a la derecha de la puntuación estándar; y para pruebas de dos colas, es el doble del valor que encontraríamos si fuésemos a calcular el valor P para una prueba de una cola.
- Si el resultado es estadísticamente significativo al nivel elegido (por lo regular 0.05 o 0.01 de nivel de significancia), rechazamos la hipótesis nula. Si no es estadísticamente significativo, no rechazamos la hipótesis nula.

EJEMPLO 1 Tasa de desempleo local

Suponga que la tasa de desempleo nacional es 3.5% y en una encuesta de $n = 450$ personas en un condado rural de Wisconsin, 22 personas aparecen como desempleadas. Las autoridades del condado hacen una solicitud de ayuda estatal con base en la afirmación de que la tasa de desempleo local es mayor que el promedio nacional. Pruebe esta afirmación a un nivel de 0.05 de significancia.

Solución Sigamos el proceso de cuatro pasos original de la página 376.

- Paso 1. La tasa de desempleo es una proporción poblacional (en contra de una media poblacional). La hipótesis nula es la suposición de que la tasa de desempleo local es igual a la tasa nacional, o $H_0: p = 0.035$. La hipótesis alternativa es la afirmación del condado de que la tasa de desempleo local es mayor que el promedio nacional, o $H_a: p > 0.035$. Observe que ésta será una prueba de cola derecha.
- Paso 2. Los estadísticos muestrales son el tamaño de la muestra, $n = 450$, y la proporción muestral, $\hat{p}$. Con base en los datos dados, la proporción muestral es

$$\hat{p} = \frac{22}{450} = 0.0489$$

- Paso 3. Ahora determinamos la posibilidad que, *bajo la suposición que la hipótesis nula sea verdadera*, se tenga oportunidad de obtener una proporción muestral al menos tan extrema como la $\hat{p} = 0.0489$ encontrada para esta muestra particular. Para determinar esta posibilidad, iniciamos por calcular la puntuación estándar para esta proporción muestral:

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} = \frac{0.0489 - 0.035}{\sqrt{0.035(1-0.035)/450}} = 1.60$$

Esta puntuación estándar es cercana al valor crítico de $z = 1.645$ para una prueba de cola derecha, pero no es mayor que este valor, por lo que el resultado no es significativo al nivel 0.05. Del apéndice A, la puntuación estándar de 1.60 tiene un valor P de $1 - 0.9452 = 0.0548$, confirmando que hay más de 0.05 de probabilidad que esta muestra surgiera por azar si la hipótesis nula es verdadera. La figura 9.8 muestra estas ideas sobre la distribución muestral.

- Paso 4. Puesto que la prueba no cumple completamente el criterio para significancia al nivel 0.05, no rechazamos la hipótesis nula. En otras palabras, esta muestra no proporciona evidencia suficiente para apoyar la afirmación de que la tasa de desempleo en el condado está por arriba del promedio nacional.

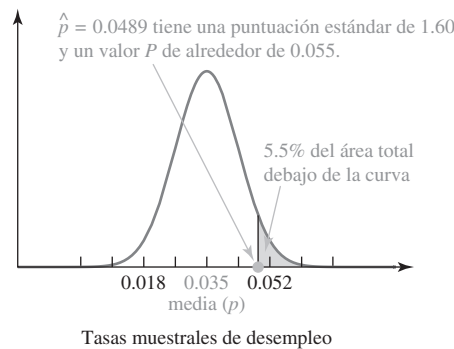


Figura 9.8 La distribución de proporciones muestrales para el ejemplo 1. La proporción muestral de $\hat{p} = 0.0489$ no está suficientemente lejos, a la derecha, para una significancia al nivel 0.05.

EJEMPLO 2 Población zurda

Una muestra aleatoria de $n = 750$ personas es seleccionada, de ellas 92 son zurdas. Utilice esta información muestral para probar la afirmación de que 10% de la población es zurda.

Solución Nuevamente seguimos el proceso de cuatro pasos.

Paso 1. La afirmación tiene que ver con la *proporción* de la población que es zurda, por lo que es una prueba con una proporción poblacional. La hipótesis nula es la afirmación de que 10% de la población es zurda, o $H_0: p = 0.1$. Para probar esta afirmación necesitamos tomar en cuenta la posibilidad de que la proporción poblacional real sea menor o mayor que 10%. Por tanto, la hipótesis alternativa es $H_a: p \neq 0.1$, que requiere una prueba de dos colas.

Paso 2. Los estadísticos muestrales son el tamaño de la muestra, $n = 750$, y la proporción de personas zurdas en la muestra:

$$\hat{p} = \frac{92}{750} = 0.123$$

Paso 3. La puntuación estándar para esta proporción muestral ($\hat{p} = 0.123$) es

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} = \frac{0.123 - 0.1}{\sqrt{0.1(1-0.1)/750}} = 2.09$$

Del recuadro de la página 386, los valores críticos para significancia al nivel 0.05 en una prueba de dos colas son puntuaciones estándar menores que -1.96 o mayores que 1.96 . La calificación estándar de 2.09 para esta prueba es mayor que 1.96 , por lo que concluimos que la prueba es significativa al nivel 0.05. Del apéndice A, el área a la derecha de una puntuación estándar de 2.09 es $1 - 0.9817 = 0.0183$. Éste sería el valor P si estuviésemos realizando una prueba de una cola. Puesto que tenemos una prueba de dos colas, lo duplicamos para determinar que el valor P es $2 \times 0.0183 = 0.0366$. La figura 9.9 muestra el significado de este valor P en la distribución muestral.

Paso 4. Puesto que la prueba es significativa al nivel 0.05, rechazamos la hipótesis nula y concluimos que la proporción de la población que es zurda *no* es igual a 10%. Recordando que el nivel de significancia es la probabilidad de un error de tipo I de rechazar una hipótesis que en realidad es verdadera.

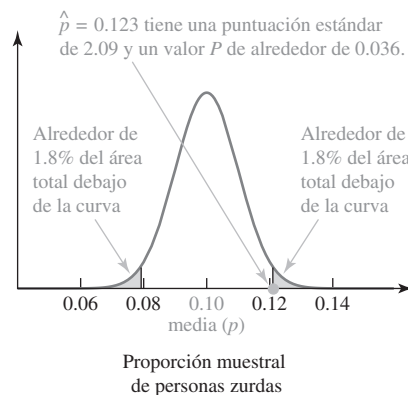


Figura 9.9 Esta gráfica muestra la posición de la proporción muestral ($\hat{p} = 0.123$) en la distribución de proporciones muestrales para el ejemplo 2. Puesto que es una prueba de dos colas, el valor P corresponde al área debajo de la curva que está a más de 2.09 desviaciones estándar, en *cualquiera* de las dos direcciones, del pico.

Sección 9.3 Ejercicios

Alfabetización estadística y pensamiento crítico

- Notación.** ¿Qué representan p , $\hat{p}$ y el valor P ?
- Distribución.** Al realizar una prueba de hipótesis como se describió en esta sección, ¿qué distribución particular se utiliza? ¿Por qué?
- Muestreo.** America OnLine realiza una encuesta en la que pide a los usuarios de internet que respondan una pregunta. Entre las 96772 respuestas, hay 76885 respuestas *sí*. ¿Es válido usar estos resultados muestrales para probar la afirmación de que la mayoría de la población general responde *sí* a esta pregunta? ¿Por qué sí o por qué no?
- Valor P .** Un valor P de 0.00001 se obtuvo al usar datos muestrales para probar la afirmación de que la mayoría de los accidentes automovilísticos ocurren a 5 millas del hogar. ¿Qué nos dice este valor P ?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Hipótesis nula.** Para probar la afirmación de que una mayoría de estadounidenses está a favor de registrar todas las armas de fuego, la hipótesis nula es $p = 0.5$.
- Hipótesis alternativa.** Para probar la afirmación de que una mayoría de estadounidenses está a favor del registro de todas las armas de fuego, la hipótesis alternativa es $p > 0.5$.
- Hipótesis alternativa.** En una prueba de hipótesis de dos colas de una afirmación acerca de una proporción, el valor P es el área a la derecha de la puntuación estándar, z .
- Hipótesis alternativa.** Se afirmó que $p > 0.5$ nunca puede respaldarse si la proporción muestral $\hat{p}$ es menor que 0.5.

Conceptos y aplicaciones

Pruebas de hipótesis. Para los ejercicios 9 al 18 utilice un nivel de significancia de 0.05 para realizar una prueba de hipótesis usando el procedimiento de cuatro pasos descritos en el texto. Asegúrese de establecer sus conclusiones.

- Encuesta de votantes.** En una encuesta previa a las elecciones, un candidato para el puesto de abogado del distrito recibe 205 de 400 votos. Suponiendo que las personas encuestadas representan una muestra aleatoria de la población de votantes, pruebe la afirmación de que una mayoría de los votantes apoya al candidato.
- Mujeres en la universidad.** De acuerdo con la Oficina de Censo de Estados Unidos, 58% de los estudiantes universitarios de 25 años o más son mujeres. Suponga que en el Colegio Clarion, en una muestra aleatoria de 2 100 estudiantes

de 25 años de edad o mayores, 1 234 son mujeres. Pruebe la afirmación de que el porcentaje de mujeres mayores en la universidad está por arriba del promedio nacional.

- Presión por calificaciones.** Un estudio solicitado por el Departamento de Educación de Estados Unidos, con base en las respuestas de 1 015 adolescentes seleccionados aleatoriamente, concluyó que 44% de ellos cita las calificaciones como su mayor fuente de presión. Pruebe la afirmación de que menos de la mitad de todos los adolescentes en la población sienten que las calificaciones son su principal fuente de presión.
- Adultos casados.** En un año reciente, 125.8 millones de adultos, o 58.6% de la población adulta de Estados Unidos, estaba casada. En un pueblo de Nueva Inglaterra, una muestra aleatoria simple de 1 445 adultos incluye 56.0% que están casados. Pruebe la afirmación de que esta muestra proviene de una población con un porcentaje de casados menor que 58.6%.
- Uso de drogas.** Un estudio del Departamento de Salud y Servicios Humanos de uso ilegal de drogas entre personas de 12 a 17 años reportó una disminución en el uso (de 11.4% en 1997) a 9.9% actualmente. Suponga que una encuesta en una gran preparatoria revela que, en una muestra aleatoria de 1 050 estudiantes, 98 reporta uso ilegal de drogas. Pruebe la afirmación de la directora de que el uso ilegal de drogas en su escuela está por debajo del promedio nacional actual.
- Pobreza.** De acuerdo con estimaciones recientes, 12.1% de las 4 342 personas en el condado de Custer, Idaho, viven en pobreza. Suponga que las personas en este condado representan una muestra aleatoria de todas las personas en Idaho. Con base en esta muestra, pruebe la afirmación de que la tasa de pobreza en Idaho es menor que la tasa nacional de 13.3%.
- Encuesta de aborto.** Una encuesta anual de estudiantes de primer año en la universidad, realizada por el Instituto de Investigación en Educación Superior en la UCLA, pregunta a aproximadamente a 276 000 estudiantes acerca de sus actitudes sobre una variedad de temas. De acuerdo con la encuesta reciente, 51% de los estudiantes de primer año creen que el aborto debe ser legal (bajó de 65% en 1990). Pruebe la afirmación de que más de la mitad de todos los estudiantes de primer año creen que el aborto debe ser legal.
- Fumar.** La tasa de fumadores para toda la población de estadounidenses adultos (mayores de 21 años) es 32% (Instituto Nacional sobre Abuso de Drogas de Estados Unidos). Suponga que para una muestra de 75 estudiantes de bellas artes mayores de 21 años, la tasa de fumadores es 35%. Utilice un nivel de significancia de 0.05 para probar la afirmación de que la tasa de fumadores para todos los estudiantes de bellas artes es mayor que el promedio nacional.

- 17. Uso de gas natural.** De acuerdo con la Administración de Información de Energía, 53.0% de los hogares en la nación usaron gas natural para calefacción en 1997. Una encuesta reciente de 3 600 hogares seleccionados aleatoriamente mostró que 54% usaron gas natural. Utilice un nivel de significancia del 0.05 para probar la afirmación de que la tasa nacional de 53.0% cambió.
- 18. Prueba clínica.** En pruebas clínicas del medicamento Lipitor, 863 pacientes fueron tratados y 19 de ellos experimentaron síntomas de gripa (con base en datos de Parke-Davis). Pruebe la afirmación de que el porcentaje de pacientes tratados con síntomas de gripa es mayor que la tasa 1.9% para pacientes que no se les da el tratamiento.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 9 en www.aw.com/bbt.

- 19. Población zurda.** Dada la afirmación de que 10% de los estadounidenses son zurdos, seleccione aleatoriamente al menos a 50 estudiantes en su escuela y determine si son zurdos. Pruebe la afirmación con una prueba de hipótesis formal.
- 20. Fumadores.** Utilice referencias de internet o bibliográficas para determinar la proporción de estadounidenses que fuman. Pruebe la afirmación de que la proporción de estudiantes en su escuela que fuma es diferente de la proporción de todos los estadounidenses. Recolecte los datos muestrales de por lo menos 50 estudiantes elegidos aleatoriamente.
- 21. Estudiantes mujeres en universidad.** Utilice internet o referencias bibliográficas para determinar la proporción de estudiantes universitarios en Estados Unidos que son mujeres. Pruebe la afirmación de que la proporción de estudiantes mujeres en su escuela es diferente de la proporción de todos los estudiantes universitarios de Estados Unidos. Recolecte datos muestrales de por lo menos 100 estudiantes elegidos de manera aleatoria.
- 22. Datos de condados.** El reporte *Statistical Abstract of the United States* y la Encuesta de Población Actual proporcionan una fuente amplia de estadísticas sociales, económicas y vitales en los niveles de condado, estado y locales. Utilice sus sitios web para comparar los datos estatales de la manera siguiente:
- Seleccione una variable de interés que incluya una proporción para un estado particular (por ejemplo, el porcentaje de personas que viven en la pobreza en Arizona).
 - Determine el valor nacional actual para esta variable (por ejemplo, la tasa de pobreza nacional).
 - Seleccione un condado particular dentro del estado y obtenga los datos correspondientes para el condado (por ejemplo, la tasa de pobreza en el condado de Pima, Arizona).
 - Suponiendo que el condado es una muestra aleatoria del estado, pruebe la afirmación de que el estado está por arriba o por debajo del nivel nacional en términos de esa variable.
 - Analice e interprete sus resultados. La prueba de hipótesis depende de la muestra (el condado) siendo una muestra aleatoria de la población (el estado). Asegúrese de analizar este factor en sus conclusiones.



EN LAS NOTICIAS



- 23. Pruebas de hipótesis en las noticias.** Encuentre una noticia o reporte de investigación que describa (quizá no de manera explícita) una prueba de hipótesis para una proporción poblacional. Adjunte el artículo y resuma el método usado.

Ejercicios de repaso del capítulo

1. En latas de coca, seleccionadas de manera aleatoria, se mide la cantidad de cola, en onzas. Los valores muestrales listados a continuación tienen una media de 12.19 onzas y una desviación estándar de 0.11 onzas. Suponga que queremos utilizar un nivel de significancia de 0.05 para probar la afirmación de que las latas de cola tienen una cantidad media de cola mayor que 12 onzas. Suponga que la población tiene una desviación estándar dada por $\sigma = 0.115$ onzas.

12.3	12.1	12.2	12.3	12.2	12.3	12.0	12.1	12.2
12.1	12.3	12.3	11.8	12.3	12.1	12.1	12.0	12.2
12.2	12.2	12.2	12.2	12.2	12.4	12.2	12.2	12.3
12.2	12.2	12.3	12.2	12.2	12.1	12.4	12.2	12.2

- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la hipótesis alternativa?
 - ¿Cuál es el valor de la puntuación estándar para la media muestral de 12.19 onzas?
 - ¿Cuál es el valor crítico?
 - ¿Cuál es el valor P ?
 - ¿Qué concluye? (Asegúrese de abordar la afirmación original de que la media es mayor que 12 onzas).
 - Describa un error de tipo I para esta prueba.
 - Describa un error de tipo II para esta prueba.
 - Determine el valor P si la prueba está modificada para probar la afirmación de que la media es *diferente* de 12 onzas (en lugar de ser mayor que 12 onzas).
2. En un estudio de fumadores que trataron de dejar de fumar con terapia de parches de nicotina, 39 estaban fumando un año después del tratamiento, y 32 no fumaban después de un año del tratamiento (con base en datos de “High Dose Nicotine Patch Therapy, de Dale, *et al.*, *Journal of the American Medical Association*, volumen 274, número 17).

Queremos usar un nivel de significancia de 0.05 para probar la afirmación de que entre los fumadores que trataron de dejar de fumar con terapia de parches de nicotina, la mayoría está fumando un año después del tratamiento:

- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la hipótesis alternativa?
 - ¿Cuál es el valor de la puntuación estándar para la proporción muestral?
 - ¿Cuál es el valor crítico?
 - ¿Cuál es el valor P ?
 - ¿Qué concluye? (Asegúrese de abordar la afirmación original de que entre los fumadores que tratan de dejar de fumar con terapia de parches de nicotina, la mayoría está fumando un año después del tratamiento).
 - Describa un error de tipo I para esta prueba.
 - Describa un error de tipo II para esta prueba.
 - Determine el valor P si la afirmación se modifica para establecer que la proporción es *igual* a 0.5?
3. Queremos probar la afirmación de que el método Clarke de selección de género es efectivo en el aumento de la posibilidad de que un bebé nuevo será niña. En una muestra aleatoria de 80 parejas que utilizaron el método Clarke, se encuentra que hay 35 niñas entre los 80 nacimientos de bebés. ¿Qué debemos concluir acerca de la afirmación? ¿Por qué *no* es necesario ir por todos los pasos de una prueba de hipótesis formal?
4. Mike Shanley quiere probar la hipótesis de que los hombres adultos tienen un peso medio mayor que 150 libras. Él encuesta a miembros adultos de su familia y obtiene 40 valores muestrales, lo que le lleva a respaldar la hipótesis alternativa de que la media poblacional es mayor que 150 libras. ¿Cuál es la falla fundamental en su procedimiento?

Cuestionario del capítulo

1. ¿Cuál es la hipótesis alternativa que resulta de afirmar que la proporción de hombres que hacen ejercicio es mayor que 0.2?
2. ¿Cuál es la hipótesis nula que es usada para probar la afirmación de que la proporción de hombres que hacen ejercicio es mayor que 0.2?
3. ¿Cuál es la hipótesis alternativa que resulta de afirmar que la proporción de criminales convictos que sirven tiempo en la prisión es igual a 0.6?
4. La afirmación de que $p > 0.75$, ¿es una prueba de cola izquierda, de cola derecha o de dos colas?
5. ¿Cuál es la hipótesis nula?
6. ¿Cuál es la hipótesis alternativa?
7. Si la prueba da como resultado un valor P de 0.2500, ¿qué concluye acerca de la afirmación dada?
8. ¿Cuáles son las dos posibles conclusiones a las que se puede llegar acerca de las hipótesis nulas?
9. ¿Cuáles son las dos posibles conclusiones a las que se puede llegar acerca de la afirmación de que se está probando?
10. Si de manera incorrecta concluye que los comediantes profesionales tienen una calificación media del CI mayor que 110 cuando su calificación media real del CI es 100, ¿cometió un error de tipo I o un error de tipo II?

En los ejercicios 5-10 supongamos que queremos usar un nivel de significancia de 0.05, lo cual aprueba la afirmación de que la puntuación media de coeficiente intelectual de los comediantes profesionales es superior a 110.

Uso de la tecnología

En la descripción de métodos para probar hipótesis acerca de una media poblacional, este capítulo considera sólo casos en los que se aplica la distribución normal. El capítulo 10 introduce métodos que usan una distribución t . Vea el procedimiento al final del capítulo 10 para esos otros métodos.

SPSS

Prueba de hipótesis para media: el SPSS no tiene un procedimiento para realizar prueba de hipótesis para una afirmación acerca de una media poblacional cuando la desviación estándar de la población es conocida. Vea la sección de tecnología al final del capítulo 10 para conocer un procedimiento que prueba una afirmación acerca de una media poblacional cuando la desviación estándar de la población no es conocida.

Pruebas de hipótesis para proporción: SPSS no tiene un procedimiento para realizar una prueba de hipótesis para una afirmación acerca de una proporción poblacional.

Excel

Cómo probar una afirmación acerca de una media poblacional: utilice el complemento Data Desk XL que es un complemento para Excel.

Primero ingrese los datos muestrales en la columna A. Si está usando Excel 2003, dé clic en **DDXL**, si está usando Excel 2007, dé clic en **Add-Ins**, luego dé clic en **DDXL**. Seleccione **Hypothesis Tests**. Bajo las opciones de tipo de función, seleccione **1 Var z Test**. Dé clic en el icono del lápiz e ingrese el rango de valores de datos, tal como

A1:A45, si tiene 45 valores listados en la columna A. Dé clic en **OK**. Siga los cuatro pasos listados en el cuadro de diálogo. Después de dar clic en **Compute** en el paso 4, obtendrá el valor P , el estadístico de prueba y la conclusión.

Cómo probar una afirmación acerca de una proporción poblacional: utilice el complemento Data Desk XL que es un complemento para Excel.

Primero ingrese el número de éxitos en la celda A1, e ingrese el número total de ensayos en la celda B1. Si está usando Excel 2003, dé clic en **DDXL**. Si está usando Excel 2007, dé clic en **Add-Ins**, luego dé clic en **DDXL**. Seleccione **Hypothesis Tests**. Bajo las opciones de tipo de función, seleccione **Summ 1 Var Prop Test** (para probar una proporción afirmada usando un resumen de datos para una variable). Dé clic en el icono del lápiz para "Num trials" e ingrese B1. Dé clic en **OK**. Siga los cuatro pasos listados en el cuadro de diálogo. Después de dar clic en **Compute** en el paso 4, obtendrá el valor P , el estadístico de prueba y la conclusión.

STATDISK

Seleccione **Analysis**, luego seleccione **Hypothesis Testing**. Para los métodos estudiados en este capítulo seleccione **Mean - One Sample** o bien **Proportion - One Sample**. Un cuadro de diálogo aparecerá. Dé clic en el cuadro en la esquina superior izquierda y seleccione el elemento que coincida con la afirmación de que será probada. Proceda a ingresar los otros elementos en el cuadro de diálogo, luego dé clic en **Evaluate**. Los resultados incluirán el estadístico de prueba y el valor P .

HABLEMOS DE SALUD Y EDUCACIÓN

¿Su educación le ayudará a vivir más tiempo?

En este capítulo nos concentramos en la prueba de hipótesis en su forma más sencilla, en la que probamos una sola afirmación acerca de una población y determinamos si es respaldada por la evidencia recolectada de una sola muestra. Como era de esperar, muchos de los problemas estadísticos más interesantes necesitan de un análisis mucho más complejo. Por ejemplo, considere preguntarse acerca de los factores que contribuyen a la longevidad (vida más larga).

Con frecuencia escuchamos acerca de factores que reducen nuestras vidas. Por ejemplo, fumar, consumo excesivo de alcohol y la obesidad, factores que tienden a hacer que la gente muera más joven. Conociendo tales hechos le puede ayudar a decidir qué evitar en su estilo de vida. Pero, ¿hay cosas que pueda hacer para aumentar su vida?

Aunque esta pregunta es mucho más difícil de responder que cualquier cosa que encontramos en este capítulo, puede abordarse a través de los mismos principios básicos. Por ejemplo, suponga que un investigador sospecha que comer avena puede hacer que viva más. Para probar esta sospecha, el investigador realiza una prueba de hipótesis; inicia con la hipótesis nula de que el consumo de avena no tiene efecto para alargar la vida, luego examina datos para buscar evidencia que apoyaría rechazar esta hipótesis nula y concluir que el consumo de avena en realidad alarga la vida. Las partes difíciles de esta investigación son recolectar los datos, por ejemplo, encontrar personas que consumen avena para ser comparadas con las personas que no lo hacen, y luego encontrar una manera de separar los efectos del consumo de avena de aquellos de la gran cantidad de otras variables que también podrían alargar la vida.

En las últimas décadas, los investigadores han identificado muchos factores que al parecer contribuyen a la longevidad. Por ejemplo, mayor salud está correlacionada con vida más larga. Correr desempeña un papel en alargar la vida. La dieta tiene numerosos efectos en la salud. Por lo general, el ejercicio es un factor positivo para alargar la vida. Sin embargo, un factor sorprendente que parece ser más importante que todos los demás son los años de educación. Entre más años esté en la escuela, más años vivirá, al menos en promedio. Las figuras 9.10 y 9.11 muestran algunos datos que apoyan esta conclusión.

¿Por qué más educación podría llevar a una vida más larga? Los investigadores han sugerido muchas razones posibles. Por ejemplo, las personas con más educación tienden a ser menos afectas a comportamientos que acortan la vida, tal como fumar o tomar en exceso, y también hacen ejercicio y otras actividades que pueden alargar la vida. Una hipótesis particularmente interesante es que obtener una educación siempre incluye alguna medida de sacrificio a corto plazo para ganar a largo plazo (tal como pagar por su educación ahora con la esperanza de obtener un trabajo mejor pagado a la larga), y este compromiso para aceptar una gratificación diferida ayuda a las personas a tomar toda clase de decisiones que contribuyan a alargar la vida. Por supuesto, nadie lo sabe a ciencia cierta, así que usted puede estar seguro que existen muchas más pruebas de hipótesis por hacerse.

Sin embargo, con base en los datos actuales, podemos sacar al menos una conclusión muy interesante directamente relevante para su lectura de este libro. Usted puede tomar una clase de estadística porque se requiere, pero también puede ayudarle a vivir más, en especial si le permitimos continuar en la escuela.





Figura 9.10 Gráfica de la esperanza de vida en comparación con años de educación para muchos países diferentes. La línea para cada país inicia a la izquierda con el promedio de esperanza de vida y años de educación en 1985 y termina a la derecha (en un punto) con esos promedios para 2000. El hecho de que casi todas las líneas estén inclinadas hacia arriba muestra que conforme los años de educación han aumentado, también ha aumentado la esperanza de vida. *Fuente: New York Times.*

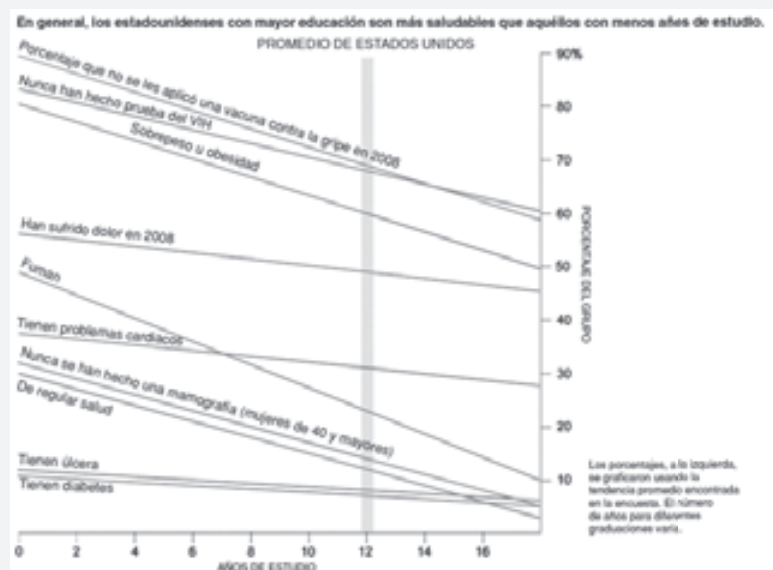


Figura 9.11 Esta gráfica muestra que todos los porcentajes de personas con una variedad de factores de riesgos de salud tienden a disminuir con el número de años de estudio en la escuela. Por ejemplo, la línea superior muestra que, en 2008, 90% de las personas con educación informal no recibieron vacuna contra la gripe, mientras que lo mismo fue cierto para sólo alrededor de 60% de personas con 16 años de educación. El hecho de que tantos factores de riesgo disminuyan con la educación, sugiere que la educación podría ser la causa común subyacente que lleva a estilos de vida más saludables. *Fuente: New York Times.*

PREGUNTAS PARA DISCUSIÓN

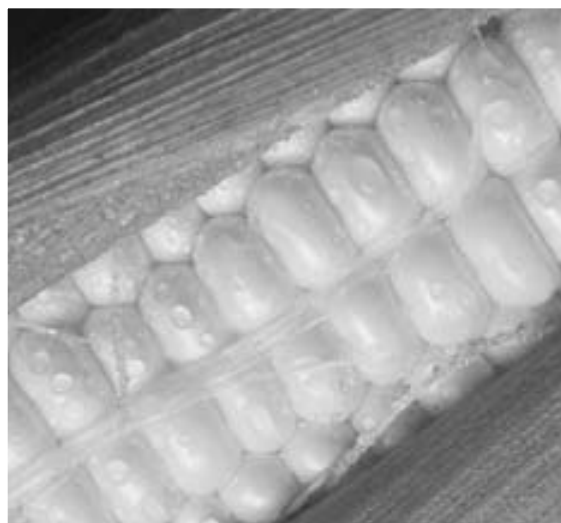
1. Las figuras 9.10 y 9.11 presentan una riqueza de datos. Interprete con cuidado cada gráfica. Asegúrese de entender el significado de cada línea (o curva). ¿Cómo cada gráfica apoya la hipótesis de que más educación conduce a una vida más larga? Explique.
2. ¿Por qué *usted* considera que la educación contribuye a una vida más larga? Por ejemplo, ¿considera que el solo hecho de estar en la escuela hace que la gente viva más? o ¿el aprendizaje extra está asociado con más años de escolaridad y esto hace la diferencia? ¿Cómo probaría su hipótesis?
3. Se sabe bien que fumar es uno de los grandes factores de riesgo en enfermedades y en muerte prematura. Como resultado, los grupos de gobierno y los de cuidado de salud con frecuencia han puesto en marcha costosas campañas de publicidad con la esperanza de convencer a la gente de dejar de fumar (¡o nunca iniciar!). Pero la figura 9.11 muestra que fumar cae dramáticamente con años de educación. Por tanto, si el objetivo es reducir fumadores, uno podría argumentar que tendría más sentido gastar dinero en programas para ayudar a la gente a permanecer en la escuela que gastar dinero en campañas antifumadores. ¿Usted cree que es buena idea? Defienda su opinión.
4. Con base en el hallazgo de que la educación contribuye a la longevidad, aunado con el hecho de que la longevidad está correlacionada con menores costos de cuidado de la salud y mayor productividad, algunas personas han propuesto que el gobierno debe aumentar los impuestos para proporcionar a más personas educación adicional. Tome una posición a favor o en contra de esta propuesta y esboce los argumentos que respalden su posición.

LECTURAS SUGERIDAS

Kolata, G., “A Surprising Secret to a Long Life: Stay in School”, *New York Times*, 3 de enero de 2007.

Lleras-Muney, Adriana, “The Relationship Between Education and Adult Mortality in the United States”, *Review of Economics Studies*, volumen 72, número 1, 2005.

HABLEMOS DE AGRICULTURA



¿Los alimentos modificados genéticamente son seguros?

Los periódicos británicos los llaman “Frankenfoods” (un juego de palabras con *Frankenstein*). Los europeos, en general, no los comen. Pero actualmente muchos alimentos vendidos en Estados Unidos ya caen en esta categoría, incluyendo un estimado de 36% de maíz estadounidense y 55% de soya estadounidense. Estamos hablando de “alimentos genéticamente modificados” o alimentos GM. Los alimentos GM son una invención de la agricultura muy reciente, que data sólo de mediados de la década de los noventa. Sin embargo, están a la vanguardia de uno de los debates más grandes en la historia de la agricultura.

Un *organismo genéticamente modificado* es un organismo en el que los científicos han insertado un gen que no existe naturalmente en el organismo o sus parientes cercanos. Los organismos genéticamente modificados están siendo probados y usados para muchos propósitos. Por ejemplo, los científicos han desarrollado bacterias que contienen genes

que producen sustancias como la insulina.

En agricultura, la modificación genética permite a los científicos crear cosechas tratadas que sería difícil o imposible de conseguir con técnicas de cultivo tradicional. Una de las primeras cosechas genéticamente modificadas y usadas con amplitud, con frecuencia llamado “maíz Bt”, ilustra la idea atrás de la modificación genética. Por lo común, el maíz es susceptible de destrucción por una variedad de plagas de insectos, lo cual hace necesario que los agricultores rocíen las cosechas con pesticidas. Por desgracia, muchos pesticidas son tóxicos para los animales, además de serlo para los insectos, y de ahí que pueden causar daños al ambiente (y podría no ser ideal para consumo humano).

Aquí es donde Bt, una bacteria conocida como *bacillus thuringiensis*, entra. La Bt vive naturalmente en la tierra. En 1911 los científicos descubrieron que la Bt produce una toxina que mata cierto tipo de insectos. Diferentes variedades de Bt matan diferentes insectos, pero por lo general son inofensivos para otros animales y los humanos. Por la década de los sesenta estas variedades han llevado al uso de Bt como un pesticida “amigable con el ambiente” que podría ser rociado en cosechas (en la forma de bacterias asesinas). Desafortunadamente, el pesticida Bt probó ser caro y con frecuencia ineficaz, principalmente porque funciona sólo si los insectos lo comen (la mayoría de los pesticidas efectivos matan al contacto). Y, puesto que se estropea rápidamente con la luz del sol y se quita de las plantas con la lluvia, mata plagas sólo si su aplicación es puesta justo a tiempo.

La modificación genética resuelve los problemas inherentes en el pesticida Bt. La acción del pesticida de bacterias Bt surge de proteínas particulares que la bacteria produce. Los científicos identificaron los genes responsables para estas proteínas, luego transfirieron estos genes a plantas tal como el maíz. Una vez que el maíz contiene los genes necesarios, produce las mismas proteínas que matan plagas como la bacteria Bt. Ya no es necesario rociar un pesticida, ahora el maíz es tóxico para las plagas. Además, como el maíz produce de manera continua las proteínas que matan plagas, ya no interesa si el pesticida se estropea o se lave con la lluvia.

La ventaja del maíz Bt sobre las cosechas tradicionales de maíz son claras, pero, ¿hay alguna desventaja? Aquí es donde el gran debate sobre los alimentos GM inicia. Por un lado, muchos científicos argumentan que los alimentos GM son completamente seguros y que sus beneficios ayudarán a mejorar la nutrición de las personas en todo el mundo. Por otra parte, están las personas, incluyendo algunos científicos que etiquetan a los alimentos GM como “Frankenfoods” y argumentan que son una de las tecnologías más peligrosas jamás inventada.

A propósito...

Más de 50 nuevas cosechas GM han sido aprobadas para la venta en Estados Unidos, incluyendo maíz y soya que produce su propio pesticida y tomates manipulados para aumentar su vida en almacenamiento. Las cosechas GM bajo desarrollo incluyen papas que resisten magulladuras, soya con mejor sabor y granos que contienen vitaminas y otros nutrientes que podrían mejorar la nutrición en países pobres.

En general, el tema de los alimentos GM seguros puede dividirse en tres preguntas principales:

1. ¿Los alimentos GM tienen algún efecto tóxico en los humanos?
2. ¿Las nuevas proteínas contenidas en los alimentos GM causan reacciones alérgicas en algunas personas?
3. ¿Las cosechas GM pueden causar algún daño imprevisto al ambiente, tal como transferir sus genes a la hierba (de este modo creando “súper hierba”) o matar animales además de plagas de insectos?

Estas preguntas pueden ser abordadas con pruebas de hipótesis. En cada caso iniciamos con una hipótesis nula: no hay diferencia en la seguridad entre alimentos tradicionales y alimentos GM. Por ejemplo, la hipótesis nula para la pregunta 1 dice que los alimentos GM no son más tóxicos que los alimentos tradicionales (que con frecuencia contienen niveles bajos de químicos tóxicos). La hipótesis alternativa establece que existe una diferencia en la toxicidad. Entonces los científicos desarrollan experimentos para probar la hipótesis. Si la evidencia proporcionada por los experimentos revela una diferencia significativa en la toxicidad entre los grupos de alimentos (más allá de la que se esperaría sólo por el azar), existe razón para rechazar la hipótesis nula.

Por desgracia, hasta la fecha los experimentos han sido incapaces de resolver la controversia. Por ejemplo, algunas cosechas genéticamente modificadas *sí* contienen productos tóxicos para los humanos y, por tanto, nunca reciben aprobación para venderse. Para los partidarios de los alimentos GM, este hecho proporciona un argumento en su favor, ya que parece mostrar que los procedimientos reguladores actuales (por ejemplo, se necesita aprobación de la Administración de Alimentos y Drogas de Estados Unidos para alimentos GM) garantizan que sólo alimentos seguros ingresan al mercado. De manera similar, mientras algunas de las proteínas en alimentos GM sin duda pueden causar reacciones alérgicas, muchos científicos piensan que ellos entienden estas reacciones lo suficientemente bien para garantizar que son aprobados alimentos seguros.

Los oponentes de los alimentos GM utilizan los mismos resultados experimentales para respaldar su caso. Argumentan que aunque los científicos hayan identificado toxicidad, obvia en los experimentos, aún pueden tener efectos a largo plazo no tomados en cuenta que podrían no aparecer en la población durante muchos años. De manera similar, aseguran que no pueden garantizar que entendamos *todas* las reacciones alérgicas y, por tanto, podrían aprobarse de forma inadvertida alimentos GM que podrían tener graves consecuencias alérgicas en al menos alguna población.

Los temas ambientales son todavía más difíciles de estudiar. Por ejemplo, un estudio ha mostrado que el maíz Bt es tóxico para la mariposa monarca, una especie que no es una plaga y que nadie quiere matar. Pero el estudio fue realizado en un laboratorio y no puede representar de manera precisa lo que ocurre en el ambiente real. De manera similar, la posibilidad de “salto de genes”, en el que genes podrían esparcirse del maíz a otras plantas, no está bien entendida y, por tanto, es muy difícil de estudiar.

Es probable que el debate sobre los alimentos GM continúe, e incluso se intensifique, durante muchos años por venir. Después de todo, cuando se trata de alimentos, todos tienen interés.

PREGUNTAS PARA DISCUSIÓN

1. Proponga un experimento que podría usar para probar la seguridad de productos alimenticios GM, tal como maíz. Describa con detalle su experimento y analice cualesquiera dificultades prácticas que podrían surgir al llevarlo a cabo o al interpretar sus resultados.
2. Ignorando cualquier tema de seguridad de alimentos GM, haga una lista de tantas formas como pueda pensar en que los alimentos GM podrían ser benéficos para la humanidad. Luego, ignorando cualquier beneficio de los alimentos GM, haga una lista de tantas formas como pueda pensar en las que los alimentos GM podrían ser peligrosos. En general, ¿considera que más uso de alimentos GM debe ser promovido o impedido? Defienda su opinión.

A propósito...

Los oponentes a los alimentos GM con frecuencia señalan al pesticida DDT como un ejemplo de cuán difícil puede ser descubrir tóxicos o efectos ambientales a largo plazo. El DDT fue usado durante 60 años antes de que los efectos perjudiciales fuesen descubiertos.

3. Algunas personas dan la elección a los consumidores requiriendo etiquetar todos los productos que contienen ingredientes genéticamente modificados. ¿Usted considera que es una buena idea? ¿Por qué sí o por qué no?
4. Investigue desarrollos recientes en el debate sobre alimentos GM. ¿La nueva información altera alguno de los argumentos principales en el debate? Explique.

LECTURAS SUGERIDAS

El debate sobre alimentos genéticamente modificados genera frecuentes encabezados. Para encontrar los últimos estudios, intente una búsqueda en la web en “GM foods” (“alimentos GM”).

Belsie, Laurent, “Superior Crops or ‘Frankenfood’?” *Christian Science Monitor*, 1 de marzo de 2000.

Barrionuevo, Alexei, “Can Gene-Altered Rice Rescue the Farm Belt?”, *New York Times*, 16 de agosto de 2005.

Rosenthal, Elisabeth, “Biotech Food Tears Rifts in Europe”, *New York Times*, 6 de junio de 2006.

Rosenthal, Elisabeth, “Questions on Biotech Crops with No Clear Answers”, *New York Times*, 6 de junio de 2006.

Yoon, Carol K., “Pollen from Genetically Altered Corn Threatens Monarch Butterfly, Study Finds”, *New York Times*, 20 mayo de 1999.



*La telaraña de este mundo
está tejida de necesidad y azar.
Sufrimiento para aquel quien se ha
acostumbrado desde su juventud
para encontrar algo necesario en
lo que es caprichoso, y para aquel
quien atribuye a algún hecho como
causa del azar y hacer una religión
de su sometimiento a dicho hecho.*

—Johann Goethe

Pruebas t , tablas de dos entradas y ANOVA

EN LOS CAPÍTULO 1 AL 9 HEMOS ESTUDIADO MUCHAS

ideas y aplicaciones troncales de estadística, si continúa su estudio de estadística encontrará muchas más aplicaciones. En este capítulo final exploramos tres aplicaciones que podría encontrar en futuros trabajos, también le ayudarán a entender el papel de la estadística en su vida diaria. Las tres aplicaciones están construidas sobre la importante técnica de prueba de hipótesis, introducida en el capítulo 9. Iniciamos con la distribución t , que se aplica a intervalos de confianza así como a pruebas de hipótesis, y luego investigamos pruebas de hipótesis con dos variables y con el método conocido como análisis de varianza o ANOVA.

OBJETIVOS DE APRENDIZAJE

10.1 La distribución t para inferencias acerca de una media

Entender cuándo es apropiado usar la distribución t de Student en lugar de la distribución normal para construir intervalos de confianza o realizar pruebas de hipótesis para medias poblacionales, y saber cómo hacer uso apropiado de la distribución t .

10.2 Pruebas de hipótesis con tablas de dos colas

Interpretar y realizar pruebas de hipótesis para independencia de variables con datos organizados en tablas de dos vías.

10.3 Análisis de varianza (ANOVA de un factor)

Interpretar y realizar pruebas de hipótesis usando el método de análisis de varianza sencillo.

10.1 La distribución *t* para inferencias acerca de una media

En la sección 8.2 analizó estimaciones de intervalos de confianza de una media poblacional. Después de hacer la suposición de que la distribución de las medias muestrales es una distribución normal, estimamos el margen de error, *E*, como

$$E \approx \frac{2s}{\sqrt{n}}$$

En una Nota técnica establecimos que la fórmula precisa para el margen de error utiliza 1.96 en lugar de 2. El valor 1.96 es la puntuación estándar *z* con un área de 0.025 a su derecha (vea el apéndice A).

La sección 9.2 analizó pruebas de hipótesis para afirmaciones acerca de una media poblacional, otra vez con base en la suposición que la distribución de muestreo es normal. Utilizamos la fórmula siguiente para la puntuación estándar de la media muestral:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

La sección 9.2 proporcionó estos criterios para rechazar la hipótesis nula (al nivel 0.05 de significancia): $z \leq -1.645$ para una prueba de cola izquierda, $z \geq 1.645$ para una prueba de cola derecha y $z \leq -1.96$ o bien $z \geq 1.96$ para una prueba de dos colas. Los valores de -1.645 , 1.645 , -1.96 y 1.96 son deducidos de la distribución normal estándar.

En aplicaciones reales, los estadísticos profesionales no siempre suponen una distribución normal para intervalos de confianza y pruebas de hipótesis. (En realidad, una revisión de artículos en revistas profesionales muestra que la distribución normal rara vez es usada). La razón principal es que, como verá en la fórmula de puntuación estándar anterior, usar la distribución normal obliga a conocer la desviación estándar de la población σ . Puesto que por lo general no conocemos σ , debemos estimarla con la desviación estándar muestral *s*. Por tanto, los estadísticos prefieren un enfoque que no necesite el conocimiento de σ . Tal es el caso con la **distribución *t* de Student**, o brevemente, **distribución *t***, que puede usarse cuando no conocemos la desviación estándar poblacional y ya sea que el tamaño de la muestra es mayor que 30 o bien la población tiene una distribución normal.

A propósito...

La distribución *t* de Student fue desarrollada por William Gosset (1876-1937), un empleado de la cervecería Guinness quien necesitaba una distribución que pudiese usarse con muestras relativamente pequeñas. La cervecería irlandesa donde trabajaba no permitía la publicación de resultados de investigación, por lo que Gosset los publicó bajo el seudónimo *Student*.

Inferencias acerca de una media poblacional: elección entre distribución *t* y distribución normal

distribución <i>t</i>	La desviación estándar poblacional no es conocida y la población se distribuye normalmente.
o	La desviación estándar poblacional no es conocida y el tamaño de la muestra es mayor que 30.
Distribución normal	La desviación estándar poblacional es conocida y la población se distribuye normalmente
o	La desviación estándar poblacional es conocida y el tamaño de la muestra es mayor que 30.

En este libro no abordaremos con detalle la naturaleza de la distribución *t*, pero la idea básica se presenta de una manera fácil de entender: la distribución *t* es muy similar en forma y simetría a la distribución normal, pero toma en cuenta la mayor variabilidad que se espera con muestras pequeñas. La figura 10.1 contrasta la distribución *t* con la distribución normal

estándar para muestras de tamaños $n = 3$ y $n = 12$. Observe que para la muestra de tamaño mayor, la distribución t es más parecida a la distribución normal. De hecho, entre mayor sea la muestra, más se parece la distribución t a la distribución normal.

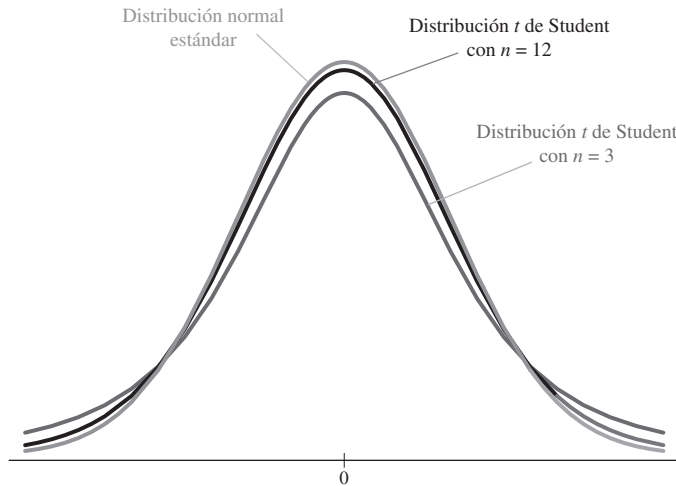


Figura 10.1 Esta figura compara la distribución normal con la distribución t para dos tamaños de muestra diferentes. Observe que conforme el tamaño de la muestra crece, la distribución t se aproxima a la distribución normal.

Entonces, el valor real de la distribución t es que nos permite extender las ideas de intervalos de confianza o pruebas de hipótesis para muchos casos en los que no podemos utilizar la distribución normal porque no conocemos la desviación estándar de la población. Sin embargo, tenga en mente que no funciona en todos los casos. Por ejemplo, si tenemos una muestra pequeña de tamaño 30 o menor o los datos sugieren que la población tiene una distribución que es radicalmente diferente de una distribución normal, entonces no se pueden utilizar ni la distribución t ni la distribución normal. Tales casos necesitan de otros métodos que no se estudian en este libro.

Intervalos de confianza usando la distribución t

De forma muy parecida a la que hicimos en las secciones 8.2 y 9.2 bajo la suposición de una distribución normal, podemos utilizar la distribución t para construir intervalos de confianza para una media poblacional o para realizar una prueba de hipótesis para una media poblacional. Sin embargo, en lugar de determinar la significancia con base en puntuaciones estándar z , tal como -1.645 , 1.645 , -1.96 o 1.96 , utilizamos valores de t , que se muestran en tablas similares a la 10.1. (Programas estadísticos también pueden generar valores de t).

Iniciemos con intervalos de confianza. Para especificar un intervalo de confianza, primero debemos calcular el margen de error, E . Con una distribución t , la fórmula es

$$E = t \cdot \frac{s}{\sqrt{n}}$$

dónde n es el tamaño de la muestra, s es la desviación estándar muestral y t es un valor que buscamos en la tabla 10.1. El único “truco” está en determinar el valor correcto de t en la tabla, que hacemos como sigue:

- Primero determine el número de **grados de libertad** para los datos muestrales, definido como el tamaño de la muestra menos uno:

$$\text{grados de libertad para distribución } t = n - 1$$

Tabla 10.1 Valores críticos de t

Grados de libertad ($n - 1$)	Área en una cola	
	0.025	0.05
	Área en dos colas	
	0.05	0.10
1	12.706	6.314
2	4.303	2.920
3	3.182	2.353
4	2.776	2.132
5	2.571	2.015
6	2.447	1.943
7	2.365	1.895
8	2.306	1.860
9	2.262	1.833
10	2.228	1.812
11	2.201	1.796
12	2.179	1.782
13	2.160	1.771
14	2.145	1.761
15	2.131	1.753
16	2.120	1.746
17	2.110	1.740
18	2.101	1.734
19	2.093	1.729
20	2.086	1.725
21	2.080	1.721
22	2.074	1.717
23	2.069	1.714
24	2.064	1.711
25	2.060	1.708
26	2.056	1.706
27	2.052	1.703
28	2.048	1.701
29	2.045	1.699
30	2.042	1.697
31	2.040	1.696
32	2.037	1.694
34	2.032	1.691
36	2.028	1.688
38	2.024	1.686
40	2.021	1.684
50	2.009	1.676
100	1.984	1.660
Mayores	1.960	1.645

- La tabla muestra grados de libertad en la columna 1. Ahora encuentre el renglón correspondiente al número de grados de libertad en sus datos muestrales y, luego, en ese renglón busque para determinar el valor apropiado de t . Para intervalos de confianza con medias poblacionales, los valores t que corresponden a 95% de confianza están en la columna 2 y en la 3 a la confianza de 90%. (Utilizamos los valores de la tabla para el “área de dos colas” ya que el margen de error puede ser por abajo o por arriba de él; por ejemplo, confianza del 95% significa que estamos buscando un área total de 0.05 alejada a ambos lados de una distribución t , como las que se muestran en la figura 10.1).

Una vez que encontró el valor t para sus datos y el nivel de confianza, puede determinar el intervalo de confianza al igual que lo hicimos en la sección 8.2, salvo que ahora se usa la nueva fórmula para el margen de error, E .

Intervalo de confianza para una media poblacional (μ) con la distribución t

Si las condiciones requieren del uso de la distribución t (σ no conocida y $n > 30$ o población distribuida normalmente), el intervalo de confianza para el valor verdadero de la media poblacional (μ) se extiende desde la media muestral menos el margen de error ($\bar{x} - E$) hasta la media muestral más el margen de error ($\bar{x} + E$). Esto es, el intervalo de confianza para la media poblacional es

$$\bar{x} - E < \mu < \bar{x} + E \text{ (o de forma equivalente, } \bar{x} \pm E)$$

donde el margen de error es

$$E = t \cdot \frac{s}{\sqrt{n}}$$

y determinamos t con base en la tabla 10.1.

EJEMPLO 1 Intervalo de confianza para presión diastólica sanguínea

A continuación se tienen cinco mediciones de presión diastólica sanguínea de adultos hombres seleccionados de manera aleatoria: 78, 54, 81, 68, 66. Estos cinco valores resultan en los estadísticos muestrales: $n = 5$, $\bar{x} = 69.4$, $s = 10.7$. Con esta muestra construya el intervalo de confianza al 95% que estime la media de la presión diastólica sanguínea para la población de todos los adultos hombres.

Solución Puesto que la desviación estándar poblacional no es conocida y es razonable suponer que los niveles de presión sanguínea de adultos hombres se distribuyen de manera normal, utilizamos la distribución t en lugar de la distribución normal. Con una muestra de tamaño $n = 5$, el número de grados de libertad es

$$\text{grados de libertad para la distribución } t = n - 1 = 5 - 1 = 4$$

Para una confianza de 95% utilizamos la columna 2 en la tabla 10.1, para encontrar que $t = 2.776$. Ahora usamos este valor junto con el tamaño de muestra dado ($n = 5$) y la desviación estándar muestral ($s = 10.7$) para calcular el margen de error, E :

$$E = t \cdot \frac{s}{\sqrt{n}} = 2.776 \cdot \frac{10.7}{\sqrt{5}} = 13.3$$

Por último, utilizamos el margen de error y la media muestral para determinar el intervalo de confianza al 95%:

$$\begin{aligned} \bar{x} - E &< \mu < \bar{x} + E \\ 69.4 - 13.3 &< \mu < 69.4 + 13.3 \\ 56.1 &< \mu < 82.7 \end{aligned}$$

Con base en las cinco mediciones muestrales, tenemos 95% de confianza de que los límites de 56.1 y 82.7 contienen la media del nivel de presión diastólica sanguínea para la población de todos los hombres adultos.

Pruebas de hipótesis usando la distribución t

Cuando la distribución t se utiliza para una prueba de hipótesis de una afirmación acerca de una media poblacional ($H_0: \mu = \text{valor afirmado}$), el valor t desempeña el papel de la puntuación estándar z cuando estudiamos estas pruebas de hipótesis en la sección 9.2. Recuerde que determinamos la significancia estadística comparando la puntuación estándar de la media muestral con el valor crítico o para determinar su valor P . Con la distribución t , en lugar de calcular la puntuación estándar z , utilizamos la fórmula siguiente para calcular t :

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

donde n es el tamaño de la muestra, $\bar{x}$ es la media muestral, s es la desviación estándar muestral y μ es la media poblacional afirmada por la hipótesis nula.

Una vez que hemos calculado t , decidimos si se rechaza o no la hipótesis nula comparando nuestro valor de t con los valores críticos de t encontrados en la tabla 10.1. Los valores críticos dependen del tipo de prueba como sigue:

Prueba de cola derecha: rechazar la hipótesis nula si el estadístico de prueba calculado t es mayor o igual al valor de t encontrado en la columna de la tabla 10.1 etiquetada como “Área de una cola”. Observe que para la prueba de una cola, la columna 2 proporciona valores críticos para la significancia al nivel 0.025 y la columna 3 proporciona valores críticos para significancia al nivel 0.05.

Prueba de cola izquierda: rechazar la hipótesis nula si el estadístico de prueba calculado t es menor o igual al negativo del valor de t encontrado en la columna de la tabla 10.1 etiquetada como “Área de una cola”. Nuevamente, puesto que es una prueba de una cola, la columna 2 proporciona valores críticos para la significancia al nivel 0.025 y la columna 3 proporciona valores críticos para significancia al nivel 0.05.

Prueba de dos colas: rechazar la hipótesis nula si el valor absoluto del estadístico de prueba calculado t es mayor o igual al valor de t encontrado en la columna de la tabla 10.1 rotulado “Área en dos colas”. Para este caso, la columna 2 proporciona los valores críticos para significancia al nivel 0.05 y la columna 3 proporciona valores críticos para significancia al nivel 0.10.

El estadístico de prueba t calculado también puede usarse para determinar el valor P ; sin embargo, por lo regular esto es hecho con ayuda de software estadístico en lugar de tablas.

EJEMPLO 2 Prueba de hipótesis de cola derecha para una media

A continuación están listadas diez puntuaciones del CI elegidas aleatoriamente de estudiantes de estadística:

111 115 118 100 106 108 110 105 113 109

Usando métodos del capítulo 4 puede confirmar que estos datos tienen los siguientes estadísticos muestrales: $n = 10$, $\bar{x} = 109.5$, $s = 5.2$. Con un nivel de significancia 0.05, pruebe la afirmación de que los estudiantes de estadística tienen una puntuación media del CI mayor que 100, que es la puntuación media del CI de la mayoría de la población.

Solución Con base en la afirmación de que la media del CI de estudiantes de estadística es mayor que 100, utilizamos la hipótesis nula $H_0: \mu = 100$ y la hipótesis alternativa $H_a: \mu > 100$. Puesto que la desviación estándar de todas las puntuaciones del CI para la población de todos los estudiantes de estadística no es conocida y ya que es razonable suponer que las calificaciones del CI de estudiantes de estadística se distribuyen normalmente, utilizamos la distribución t en lugar de la distribución normal. El valor del estadístico t es calculado como sigue:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{109.5 - 100}{5.2 / \sqrt{10}} = 5.777$$

Ahora necesitamos comparar este valor para el valor crítico apropiado con base en la tabla 10.1:

- Encontramos el renglón correcto reconociendo que este conjunto de datos tiene $n - 1 = 10 - 1 = 9$ grados de libertad.
- Puesto que es una prueba de una cola y estamos preguntando por una prueba con significancia en el nivel 0.05, utilizamos los valores de la columna 3.
- Buscando en la fila para 9 grados de libertad y la columna 3, encontramos que el valor crítico para significancia al nivel 0.05 es $t = 1.833$.

Puesto que el estadístico muestral $t = 5.777$ es mayor que el valor crítico $t = 1.833$, rechazamos la hipótesis nula. Concluimos que hay evidencia suficiente para respaldar la afirmación de que la puntuación media del CI es mayor que 100.

Podemos ser más precisos usando software para calcular el valor P de esta prueba de hipótesis, que resulta ser 0.000135. Observe que este valor P es mucho menor que 0.05, por lo que podemos estar muy confiados en la decisión de rechazar la hipótesis nula y respaldar la afirmación de que la puntuación media del CI es mayor que 100.

EJEMPLO 3 Prueba de hipótesis de dos colas para una media

Con los datos muestrales del ejemplo 2 y el mismo nivel de significancia de 0.05, pruebe la afirmación de que la puntuación media del CI de estudiantes de estadística es *igual a* 100.

Solución Con base en la afirmación de que la puntuación media del CI de estudiantes de estadística es igual a 100, utilizamos la hipótesis nula $H_0: \mu = 100$ y la hipótesis alternativa $H_a: \mu \neq 100$. La forma “diferente a” de la hipótesis alternativa significa que es una prueba de dos colas.

El valor del estadístico muestral sigue siendo el mismo que el del ejemplo 2 ($t = 5.777$), pero los valores críticos son diferentes para la prueba de dos colas. Puesto que estamos usando el mismo conjunto de datos, el número de grados de libertad aún es 9, que nos dice cuál renglón buscar en la tabla 10.1. Pero para una prueba de dos colas encontramos el valor crítico para la significancia al nivel 0.05 en la columna 2, en lugar de la columna 3. El valor crítico en el renglón 9 y columna 2 es $t = 2.262$.

Observe que el valor absoluto del estadístico de prueba $t = 5.777$ es mayor que el valor crítico $t = 2.262$, así que nuevamente rechazamos la hipótesis nula. Concluimos que hay evidencia suficiente para rechazar la afirmación de que la puntuación media del CI de estudiantes de estadística es igual a 100. (Con software encontramos que el estadístico de prueba tiene un valor P de 0.000269).

EJEMPLO 4 Prueba de hipótesis de cola izquierda para una media

A consecuencia de los gastos involucrados, las pruebas de accidentes automovilísticos con frecuencia usan muestras pequeñas. En un estudio, cinco automóviles BMW son chocados bajo condiciones estándar, y los costos de reparación (en dólares) son usados para probar la afirmación de que el costo medio de reparación para todos los automóviles BMW es menor que \$3 000. La muestra de cinco costos de reparación tiene una media de \$2 835 con una desviación estándar de \$883. Utilice un nivel de significancia 0.05 para probar la afirmación de que la población de costos de reparación de BMW tiene una media menor que \$3 000.

Solución Con base en la afirmación de que la media del costo de reparación es menor que \$3 000, la hipótesis nula es $H_0: \mu = \$3\,000$ y la hipótesis alternativa es $H_a: \mu < \$3\,000$. La desviación estándar de la población no es conocida. Supondremos que los costos de reparación se distribuyen normalmente (la prueba de hipótesis para t no es muy sensible a pequeñas desviaciones de las distribuciones normales, por lo que se trata de una suposición razonable). El valor del estadístico de prueba t es

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{2835 - 3000}{883 / \sqrt{5}} = -0.418$$

Para este ejemplo, el tamaño de la muestra es $n = 5$, por lo que el número de grados de libertad es $n - 1 = 5 - 1 = 4$. Ya que es una prueba de una cola, encontramos los valores críticos para significancia al nivel 0.05 en la columna 3. Si revisa en el renglón 4 y columna 3, encontrará el valor crítico $t = 2.132$. Para una prueba de cola *izquierda* estamos buscando valores menores o iguales al negativo de este valor, o $t = -2.132$.

Puesto que el estadístico de prueba muestral $t = -0.418$ no es menor o igual al valor crítico $t = -2.132$, no rechazamos la hipótesis nula. No hay evidencia suficiente para respaldar la afirmación de que el costo medio de reparación sea menor que \$3 000. Con software encontraría que el valor P para el estadístico de prueba es 0.3488. En otras palabras, si la hipótesis nula es verdadera, la probabilidad de seleccionar una muestra al menos tan extrema como la encontrada aquí es alrededor de 35%. Ésta es una probabilidad muy alta que no nos da razón para pensar que la hipótesis nula sea errónea.

Sección 10.1 Ejercicios

Alfabetización estadística y pensamiento crítico

- Distribución t .** Suponga que quiere construir un intervalo de confianza o probar una hipótesis acerca de una media poblacional. Describa las condiciones que indican que debe usar la distribución t .
- Distribución t .** Al construir un intervalo de confianza para estimar una media poblacional o probar una hipótesis acerca de una media poblacional, ¿por qué la distribución t es usada con mucha mayor frecuencia que la distribución normal?
- Terminología.** ¿Cuál es la diferencia entre una distribución t de Student y una distribución t ?
- Método de muestreo.** Suponga que quiere estimar la cantidad media de efectivo que llevan las personas en Estados Unidos. Si encuesta a 12 de sus mejores amigos, ¿puede utilizar la distribución t para desarrollar un intervalo de confianza? ¿El resultado sería bueno para estimar la media poblacional?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

- Prueba t .** Al probar la hipótesis acerca de una media poblacional con una muestra que tiene menos de 30 valores, la distribución t siempre es usada.
- Método de muestreo.** Al probar una hipótesis acerca de una media poblacional con una muestra que parezca provenir de una población distribuida normalmente con una desviación estándar poblacional desconocida, la distribución t puede usarse sin importar el método que fue usado para recolectar los datos muestrales.
- Prueba t .** Cuando se usa la distribución t para probar la afirmación de que los estudiantes tienen una puntuación media del CI mayor que 105, si la media muestral es 103, la

prueba de hipótesis no es necesaria y podemos concluir que no existen datos suficientes para respaldar la afirmación.

- Distribución t frente a distribución normal.** Puesto que las pruebas t no necesitan un valor conocido de la desviación estándar poblacional y ya que el valor de la desviación estándar poblacional raramente es conocido, la prueba t es usada en situaciones reales con mucha más frecuencia que la prueba de hipótesis que usa la distribución normal.

Conceptos y aplicaciones

Intervalos de confianza. En los ejercicios 9 al 16 usamos la distribución t para construir el intervalo de confianza para estimar la media poblacional.

- Puntuaciones del CI.** Una muestra aleatoria simple de puntuaciones del CI se selecciona de una población distribuida normalmente. Los estadísticos muestrales son $n = 25$, $\bar{x} = 102.5$, $s = 12.8$. Construya el intervalo de confianza al 95% para estimar la media de la población.
- Estaturas de jugadores de la NBA.** En una muestra aleatoria simple de estaturas de jugadores de baloncesto en la NBA la población tiene una distribución que es aproximadamente normal. Los estadísticos muestrales son $n = 16$, $\bar{x} = 77.9$ pulgadas, $s = 3.50$ pulgadas. Construya el intervalo de confianza al 95% para estimar la media poblacional.
- Longitud del codo a la punta del dedo de hombres.** Se obtiene una muestra aleatoria simple de hombres y se mide la longitud desde el codo hasta la punta del dedo. La población de esas longitudes tiene una distribución que es normal. Los estadísticos muestrales son $n = 35$, $\bar{x} = 14.5$ pulgadas, $s = 0.7$ pulgadas. Construya el intervalo de confianza al 95% para estimar la media poblacional.
- Calificaciones del SAT.** Una muestra aleatoria simple de calificaciones del SAT se obtiene y la población tiene una distribución que es aproximadamente normal. Los estadísticos muestrales son $n = 41$, $\bar{x} = 1503$, $s = 352$. Construya el intervalo de confianza al 95% para estimar la media poblacional.

13. Costos hospitalarios de accidentes. Se realizó un estudio para estimar los costos de hospitalización para víctimas de accidentes que utilizaron cinturones de seguridad. Veinte casos, elegidos aleatoriamente, tienen una distribución que parece tener aproximadamente una forma de campana con una media de \$9004 y una desviación estándar de \$5629 (con base en datos del Departamento de Transporte de Estados Unidos).

- Construya el intervalo de confianza al 95% para la media de tales costos.
- Si usted es un gerente de una compañía de seguros que proporciona las tasas más bajas para conductores que usan cinturones de seguridad y quiere una estimación conservadora para el peor de los escenarios, ¿qué cantidad debe usar como costos hospitalarios posibles para la víctima de un accidente que usa un cinturón de seguridad?

14. Temperaturas pronosticadas y temperaturas reales. Uno de los autores compiló una lista de temperaturas máximas reales y la correspondiente lista de temperaturas máximas pronosticadas con tres días de anticipación. Luego se determinó la diferencia de cada día restando la temperatura pronosticada de la temperatura real; el resultado fue una lista de 31 valores con una media de -0.419° y una desviación estándar de 3.704° .

- Construya el intervalo de confianza al 95% para estimar la diferencia media entre todas las temperaturas máximas reales y las temperaturas máximas pronosticadas con tres días de anticipación.
- ¿El intervalo de confianza incluye a 0° ? Si el intervalo de confianza incluye a 0° , ¿podemos concluir que las temperaturas pronosticadas con tres días de anticipación son imprecisas?

15. Estimación de contaminación de automóviles. Cada automóvil en una muestra de siete automóviles fue probado para emisiones de óxido de nitrógeno (en gramos por milla) y se obtuvieron los resultados siguientes: 0.06, 0.11, 0.16, 0.15, 0.14, 0.08, 0.15 (con base en datos de la Agencia de Protección al Ambiente).

- Suponiendo que esta muestra es representativa de todos los automóviles en uso, construya el intervalo de confianza al 95% para estimar la cantidad media de emisiones de óxido de nitrógeno para todos los automóviles.
- La Agencia de Protección al Ambiente requiere que las emisiones de óxido de nitrógeno sean menores que 0.165 gramos/milla. Alguien afirmó que las emisiones de óxido de nitrógeno tienen una media igual a 0.165 gramos/milla. ¿El intervalo de confianza sugiere que esta afirmación no es válida? ¿Por qué sí o por qué no?

16. Monitoreo de plomo en el aire. A continuación están listadas las cantidades de plomo medidas (en microgramos por metro cúbico, o $\mu\text{g}/\text{m}^3$) en el aire.

5.40 1.10 0.42 0.73 0.48 1.10

La Agencia de Protección al Ambiente ha establecido una calidad estándar del aire para plomo de $1.5 \mu\text{g}/\text{m}^3$. Las

mediciones fueron registradas en el sitio del edificio 5 del World Trade Center en diferentes días inmediatamente después de la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Después del colapso de los dos edificios del World Trade Center, había considerable interés por la calidad del aire. Utilice los valores dados para construir un intervalo de confianza al 95% para estimar la cantidad media de plomo en el aire. ¿Hay algo en el conjunto de datos que sugiera que el intervalo de confianza podría no ser bueno? Explique.

Pruebas de hipótesis. En los ejercicios 17 al 24 pruebe la afirmación dada.

17. Azúcar en el cereal. Una muestra aleatoria simple de 16 cereales diferentes se obtuvo, y para cada cereal seleccionado se midió el contenido de azúcar (en gramos de azúcar por gramo de cereal). Esas cantidades tienen una media de 0.295 gramos y una desviación estándar de 0.168 gramos. Utilice un intervalo de confianza de 0.05% para probar la afirmación de un promotor de cereales de que la media de todos los cereales es menor que 0.3 gramos.

18. Prueba de precisión de los relojes de pulsera. Se seleccionaron aleatoriamente estudiantes y se midió la precisión de sus relojes de pulsera, con errores positivos representando a relojes que estaban adelantados a la hora correcta y errores negativos representando relojes que estaban atrasados de la hora correcta. Los 30 valores tienen una media de 117.3 segundos y una desviación estándar de 185.0 segundos. Utilice un nivel de significancia de 0.05 para probar la afirmación de que la población de todos los relojes tiene una media igual a 0 segundos. ¿Qué puede concluir acerca de la precisión de los relojes de pulsera de la gente?

19. Bolas de béisbol. En pruebas anteriores, bolas de béisbol se dejaron caer 24 pies sobre una superficie de concreto y rebotaron un promedio de 92.84 pulgadas. En una prueba de una muestra de 20 bolas nuevas, las alturas de los rebotes tuvieron una media de 92.67 pulgadas y una desviación estándar de 1.79 pulgadas (con base en datos del Laboratorio Nacional Brookhaven y *USA Today*). Utilice un nivel de significancia de 0.05 para determinar si hay evidencia suficiente para respaldar la afirmación de que las nuevas bolas de béisbol tienen alturas de rebote con una media diferente de las 92.84 pulgadas. ¿Parece que las nuevas bolas de béisbol son diferentes?

20. Confiabilidad de radios en aeronaves. El tiempo medio entre fallas para un radio de la compañía Telektronic usado en aeronaves ligeras es 420 horas. Después que 15 radios nuevos fueron modificados en un intento por mejorar la confiabilidad, se realizaron pruebas para medir los tiempos entre fallas. Los 15 radios tuvieron un tiempo medio entre fallas de 442 horas con una desviación estándar de 44.0 horas. Utilice un nivel de significancia del 0.05 para probar la afirmación de que los radios modificados tienen un tiempo medio entre fallas que es mayor a 420 horas. ¿Parece que la modificación mejoró la confiabilidad?

21. Efecto de complemento vitamínico en el peso al nacer.

Cuando fueron registrados los pesos al nacer de una muestra aleatoria simple de 16 bebés hombres de madres que tomaron un complemento vitamínico especial, la muestra tuvo una media de 3.675 kilogramos y una desviación estándar de 0.657 kilogramos (con base en datos del Departamento de Salud del Estado de Nueva York). Utilice un nivel de significancia de 0.05 para probar la afirmación de que el peso medio al nacer para todos los bebés varones de madres que tomaron vitaminas es diferente de 3.39 kilogramos, que es la media para la población de todos los hombres. Con base en estos resultados, ¿el complemento de vitaminas parece tener un efecto en el peso al nacer?

- 22. Pulso.** Uno de los autores afirmó que su pulso cardíaco era menor que el pulso medio de los estudiantes de estadística. El pulso del autor fue medido y se encontró que era de 60 latidos por minuto, y los 20 estudiantes de su clase midieron sus pulsos. Los 20 estudiantes tuvieron un pulso medio de 74.5 latidos por minuto y su desviación estándar fue 10.0 latidos por minuto. ¿Hay evidencia suficiente para respaldar la afirmación de que el pulso medio de estudiantes de estadística es mayor de 60 latidos por minuto? Utilice un nivel de significancia de 0.05.

- 23. Ganadores olímpicos.** A continuación se listan los tiempos ganadores (en segundos) de hombres en 100 metros libres para juegos olímpicos consecutivos, listados en orden por renglón.

12.0	11.0	11.0	11.2	10.8	10.8	10.8	10.6
10.8	10.3	10.3	10.3	10.4	10.5	10.2	10.0
9.95	10.14	10.06	10.25	9.99	9.92	9.96	

Suponiendo que estos resultados son datos muestrales aleatoriamente seleccionados de la población de todos los juegos olímpicos pasados y futuros, utilice un nivel de significancia de 0.05 para probar la afirmación de que el tiempo medio es menor a 10.5 segundos. ¿Cuál es su ob-

servación acerca de la precisión de los números? ¿Qué características extremadamente importantes del conjunto de datos no es considerada en la prueba de hipótesis?

- 24. Nicotina en cigarros.** La compañía de tabaco Carolina anunció que sus cigarros sin filtro mejor vendidos contienen 40 miligramos de nicotina o menos, pero la revista *Consumer Advocate* realizó pruebas de 10 cigarros seleccionados aleatoriamente y encontró las cantidades (en miligramos) que se muestran a continuación.

47.3	39.3	40.3	38.3	46.3
43.3	42.3	49.3	40.3	46.3

Utilice un nivel de significancia de 0.05 para probar la afirmación de la revista de que la media de contenido de nicotina es mayor a 40 miligramos.



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 10 en www.aw.com/bbt.

- 25. Personalidades destacadas.** El *World Almanac and Book of Facts* incluye una sección denominada “Noted Personalities”, con subsecciones para arquitectos, artistas, líderes de negocios, humoristas y otras categorías. Seleccione una muestra de un grupo y encuentre la media y la desviación estándar de longevidad. Pruebe la afirmación de que el grupo tiene una longevidad diferente de 77 años, que es la longevidad media actual para la población general.

- 26. Gosset.** La distribución t fue desarrollada originalmente por William Sealy Gosset. Utilice internet para buscar “William Sealey Gosset” o “William S. Gosset”. Redacte un párrafo que describa información importante acerca de Gosset y sus logros.

10.2 Pruebas de hipótesis con tablas de dos colas

Todas las pruebas de hipótesis que hemos considerado hasta ahora han tenido hipótesis nulas en las que una media poblacional (μ) o una proporción poblacional (p) se afirma que son *iguales* a algún valor. Pero existen muchas situaciones en las que la hipótesis nula toma una forma diferente. En esta sección examinamos pruebas de hipótesis diseñadas para buscar una relación entre dos variables. El proceso básico es el mismo de siempre: identificar la hipótesis nula y la hipótesis alternativa, probar la hipótesis nula con datos muestrales, y luego decidir si la evidencia de la muestra respalda rechazar o no rechazar la hipótesis nula. Si rechazamos la hipótesis nula, significa que aceptamos la hipótesis alternativa.

Identificación de las hipótesis con dos variables

Suponga que a los administradores de un colegio les preocupa que pueda haber sesgo en la forma en que los títulos son otorgados a hombres y mujeres en diferentes departamentos. Por tanto, recolectan información sobre el número de títulos otorgados a hombres y mujeres en diferentes departamentos. Estos datos tratan con dos variables: *especialidad* y *género*. La

variable *especialidad* puede tomar diversos valores, tales como biología, negocios, matemáticas y música. La variable *género* puede tomar sólo dos valores, masculino o femenino. Para probar si existe sesgo en el otorgamiento de títulos, los administradores hacen la pregunta siguiente:

¿Los datos sugieren una relación entre las dos variables?

Si existe una relación, entonces los hombres y las mujeres son elegidos para especialidades a diferentes tasas, sugiriendo que el género de una persona influye en su elección de especialidad (ya sea por elección o a consecuencia del sesgo dentro de diferentes departamentos). Si no hay relación, significa que no hay evidencia de que el género influye en la elección de la especialidad de una persona.

Esta idea sugiere las elecciones siguientes para las dos hipótesis. La hipótesis nula, H_0 , establece que las dos variables son *independientes* (*no hay relación* entre ellas); en nuestro ejemplo actual establece que no existe relación entre género y especialidad. La hipótesis alternativa, H_a , establece lo opuesto. Hay una relación entre las variables, que en este caso implica que el género influye la elección de especialidad de una persona.

Hipótesis nula e hipótesis alternativa con dos variables

La **hipótesis nula**, H_0 , establece que las variables son independientes (*no hay relación* entre ellas).

La **hipótesis alternativa**, H_a , establece que *sí hay* una relación entre las dos variables.

Cómo mostrar los datos en tablas de dos entradas

Con la hipótesis identificada, el paso siguiente en la prueba de hipótesis es examinar el conjunto de datos para ver si respalda rechazar o no rechazar la hipótesis nula. La recolección de datos para nuestro ejemplo significa encontrar los números de hombres y mujeres con títulos otorgados en diferentes especialidades. Una vez que hemos recolectado los datos, necesitamos encontrar una manera eficiente de mostrarlo. Puesto que estamos tratando con dos variables, podemos exhibir los datos de manera eficiente con una **tabla de dos entradas** (también denominada **tabla de contingencia**), así nombrada porque exhibe dos variables.

La tabla 10.2 muestra que la tabla de dos entradas podría verse como para datos en las variables *especialidad* y *género*. Cada celda muestra una frecuencia (o cuenta) para una combinación de dos variables. Por ejemplo, la celda en el renglón *Mujeres* y columna *Biología* muestra que 32 grados de licenciatura fueron otorgados a mujeres en biología. De manera análoga, la celda en el renglón *Hombres* y columna *Administración* muestra que 87 grados de licenciatura fueron otorgados a hombres en administración.

Tabla 10.2 Tabla de dos entradas para las variables *especialidad* y *género*

	Biología	Administración	Matemáticas	Psicología	...
Mujeres	32	110	18	75	...
Hombres	21	87	15	70	...

variable 1 *especialidad* →

↑
variable 2 *género*

Nota: una variable se muestra a lo largo de las columnas y la otra a lo largo de los renglones. Aquí sólo hay dos renglones, ya que el género sólo puede ser hombre o mujer. Existen tantas columnas para las especialidades, aquí sólo se muestran las primeras.

Tablas de dos entradas

Una **tabla de dos entradas** muestra la relación entre dos variables listando una variable en los renglones y la otra variable en las columnas. Las entradas en las celdas de la tabla se denominan *frecuencias* (o *conteos*).

No puedo hacerlo sin
contadores.

—William Shakespeare,
The Winter's Tale

Si estamos buscando *cualquier* relación entre especialidad y género, necesitaríamos un conjunto completo de datos para todas las especialidades, lo que significa que la tabla 10.2 tendría docenas de columnas. Además, como veremos en breve, realizar los cálculos para la prueba de hipótesis requiere que encontremos los totales de todos los renglones y columnas. Aquí, para simplificar los cálculos, nos concentraremos sólo en dos especialidades, biología y administración. Esto es, en lugar de preguntar si hay una relación entre especialidad y género para todas las especialidades, consideramos una pregunta más sencilla: ¿El género de una persona influye en su elección de la especialidad en biología o en administración? La tabla 10.3 muestra los datos de biología y administración extraídos de la tabla 10.2, junto con los totales por renglón y por columna.

Tabla 10.3 Tabla de dos entradas para títulos en biología y en administración			
	Biología	Administración	Total
Mujeres	32	110	142
Hombres	21	87	108
Total	53	197	250

UN MOMENTO DE REFLEXIÓN

Utilice la tabla 10.3 para responder a las preguntas siguientes: (a) ¿Cuántos títulos en administración fueron concedidos a hombres? (b) ¿Cuántos títulos en administración fueron otorgados en total? (c) Compare el número total de títulos concedidos a hombres y a mujeres con el número total de títulos otorgados en administración y biología. ¿Estos totales son iguales o diferentes? ¿Por qué?

A propósito...

En todos los colegios y universidades estadounidenses, los hombres exceden a las mujeres en decidirse por especialidades de ingeniería, ciencias de la computación y arquitectura. Las mujeres exceden a los hombres en psicología, bellas artes, contabilidad, biología y educación elemental.

EJEMPLO 1 Una tabla de dos entradas para una encuesta

La tabla 10.4 muestra los resultados de una encuesta previa a la elección sobre el control de armas. Utilice la tabla para responder las preguntas siguientes:

Tabla 10.4 Tabla de dos entradas para encuesta de control de armas (con totales)				
	A favor de leyes más estrictas	En contra de leyes más estrictas	Indeciso	Total
Demócratas	456	123	43	622
Republicano	332	446	21	799
Total	788	569	64	1 421

- a. Identifique las dos variables que se muestran en la tabla.
- b. ¿Qué porcentaje de demócratas están a favor de leyes más estrictas?
- c. ¿Qué porcentaje de todos los votantes están a favor de leyes más estrictas?
- d. ¿Qué porcentaje de aquellos que se oponen a leyes más estrictas son republicanos?

Solución Observe que la suma de los totales de los renglones y la suma de los totales de las columnas son iguales.

- a. Los renglones muestran la variable *respuesta a la encuesta*, que puede ser “a favor de leyes más estrictas”, “en contra de leyes más estrictas” o “indeciso”. Las columnas muestran la variable *afiliación partidista*, que en esta tabla puede ser demócrata o republicano.
- b. De los 622 demócratas encuestados, 456 están a favor de leyes más estrictas. El porcentaje de demócratas a favor de leyes más estrictas es $456/622 = 0.733$, o 73.3%.
- c. De las 1 421 personas encuestadas, 788 están a favor de leyes más estrictas. El porcentaje de todos los encuestados a favor de leyes más estrictas es $788/1\,421 = 0.555$, o 55.5%.
- d. De las 569 personas encuestadas que se oponen a leyes más estrictas, 446 son republicanos. Como $446/569 = 0.783$, entonces 78.3% de los que se oponen a leyes más estrictas son republicanos.

Cómo llevar a cabo la prueba de hipótesis

Con la hipótesis identificada y los datos organizados en la tabla de dos entradas, ahora estamos preparados para realizar la prueba de hipótesis. Otra vez, la idea básica de la prueba de hipótesis es la misma de siempre, decidir si la información proporciona evidencia suficiente para rechazar la hipótesis nula. Para el caso de una prueba con una tabla de dos entradas, los pasos específicos son como sigue:

- Como siempre, iniciamos suponiendo que la hipótesis nula es verdadera, lo que significa que no hay relación entre las dos variables. En ese caso esperaríamos que las frecuencias (el número de individuos en las celdas individuales) en la tabla de dos entradas sea la que ocurriría por puro azar. Entonces, nuestro primer paso es encontrar una forma para calcular las frecuencias que esperaríamos por el azar.
- Luego comparamos las frecuencias esperadas por azar con las frecuencias observadas en la muestra, que son las frecuencias mostradas en la tabla. Hacemos esto por medio del cálculo de algo llamado el *estadístico ji cuadrada* para los datos muestrales, la cual aquí desempeña un papel similar al de la puntuación estándar z en las pruebas de hipótesis que realizamos en el capítulo 9 o el papel de la t en la sección 10.1.
- Recuerde que para las pruebas de hipótesis en el capítulo 9 tomamos la decisión acerca de rechazar o no rechazar la hipótesis nula al comparar el valor calculado de la puntuación estándar para la muestra con los valores críticos dados en las tablas; de manera análoga, en la sección 10.1 comparamos los valores calculados del estadístico de prueba t con los valores encontrados en una tabla. Aquí hacemos lo mismo, salvo que en lugar de usar los valores críticos para la puntuación estándar z o t , utilizamos los valores críticos para el estadístico ji cuadrada.

Como un ejemplo del proceso, desarrollamos estos pasos con los datos de la tabla 10.3.

Determinación de las frecuencias esperadas por el azar

Nuestro primer paso es encontrar las frecuencias esperadas en la tabla 10.3 como *si no hubiese relación* entre las variables, que es equivalente a la frecuencia esperada sólo por azar. Iniciamos por determinar la frecuencia que esperaríamos por azar para especialidad en administración de hombres. Para hacer esto, primero calculamos la fracción de *todos* los estudiantes en la muestra que recibieron título de administración:

$$\frac{\text{títulos totales en administración}}{\text{títulos totales}} = \frac{197}{250}$$

Como se estudió en el capítulo 6, podemos interpretar este resultado como una *probabilidad de frecuencia relativa*. Esto es, si seleccionamos a un estudiante *al azar* de la muestra, la probabilidad de que tenga un título en administración es $197/250$. Usando la notación para probabilidad, escribimos

$$P(\text{administración}) = \frac{197}{250}$$

De manera similar, si seleccionamos al azar a un estudiante en la muestra, la probabilidad que este estudiante sea hombre es

$$P(\text{hombre}) = \frac{\text{total de hombres}}{\text{total de hombres y mujeres}} = \frac{108}{250}$$

Ahora tenemos toda la información necesaria para determinar la frecuencia que esperaríamos por azar para hombres con título en administración. Recuerde de la sección 6.5 que si dos eventos A y B son independientes (el resultado de uno no afecta la probabilidad del otro), entonces

$$P(A \text{ y } B) = P(A) \times P(B)$$

Podemos aplicar esta regla para determinar la probabilidad de que un estudiante sea *a la vez* hombre y con especialidad en administración (suponiendo la hipótesis nula de que género es independiente de especialidad):

$$P(\text{hombre y administración}) = P(\text{hombre}) \times P(\text{administración}) = \frac{108}{250} \times \frac{197}{250} \approx 0.3404$$

Esta probabilidad es equivalente a la fracción del total de estudiantes que esperamos sean hombres con especialidad en administración *si no hay relación* entre género y especialidad. Por tanto, multiplicamos esta probabilidad por el número total de estudiantes en la muestra (250) para determinar el número (o frecuencia) de hombres con especialidad en administración que esperamos por azar:

$$\frac{108}{250} \times \frac{197}{250} \times 250 \approx 85.104$$

A este valor le llamamos **frecuencia esperada** para el número de hombres con especialidad en administración. (Observe la semejanza entre la idea de frecuencia esperada y el de *valor esperado* analizado en la sección 6.3).

Definición

Las **frecuencias esperadas** en una tabla de dos entradas son las frecuencias que esperaríamos por el azar *si no hubiese relación* entre las variables de los renglones y las columnas.

EJEMPLO 2 Frecuencias esperadas para la tabla 10.3

Determine las frecuencias esperadas por el azar para las tres celdas restantes en la tabla 10.3. Luego construya una tabla que muestre las frecuencias observadas y las frecuencias esperadas por el azar.

Solución Seguimos el procedimiento usado para determinar el número esperado de hombres con especialidad en administración. Ya tenemos $P(\text{hombre})$ y $P(\text{administración})$. También necesitaremos $P(\text{mujer})$ y $P(\text{biología})$:

$$P(\text{mujer}) = \frac{\text{total de mujeres}}{\text{total de hombres y mujeres}} = \frac{142}{250} = 0.5680$$

$$P(\text{biología}) = \frac{\text{total de títulos en biología}}{\text{total de títulos}} = \frac{53}{250} = 0.2120$$

Combinamos las probabilidades individuales para encontrar la probabilidad para cada una de las tres celdas restantes:

$$P(\text{mujer y administración}) = P(\text{mujer}) \times P(\text{administración}) = \frac{142}{250} \times \frac{197}{250} \approx 0.4476$$

$$P(\text{hombre y biología}) = P(\text{hombre}) \times P(\text{biología}) = \frac{108}{250} \times \frac{53}{250} \approx 0.0916$$

$$P(\text{mujer y biología}) = P(\text{mujer}) \times P(\text{biología}) = \frac{142}{250} \times \frac{53}{250} \approx 0.1204$$

Observe que, como esperábamos, el total de las probabilidades para las cuatro celdas es $0.3404 + 0.4476 + 0.0916 + 0.1204 = 1.0000$.

Ahora encontramos las frecuencias esperadas multiplicando las probabilidades de la celda por el número total de estudiantes (250):

$$\begin{array}{l} \text{Frecuencia esperada de mujeres} \\ \text{con especialidad en administración} \end{array} = 250 \times \frac{142}{250} \times \frac{197}{250} \approx 111.896$$

Millones vieron la manzana caer, pero Newton fue el único que se preguntó por qué.

—Bernard Baruch

$$\begin{aligned}\text{Frecuencia esperada para hombres con especialidad en biología} &= 250 \times \frac{108}{250} \times \frac{53}{250} \approx 22.896 \\ \text{Frecuencia esperada para mujeres con especialidad en biología} &= 250 \times \frac{142}{250} \times \frac{53}{250} \approx 30.104\end{aligned}$$

La tabla 10.5 repite la información de la tabla 10.3, pero esta vez también muestra la frecuencia esperada para cada celda (entre paréntesis). Para comprobar que hicimos el trabajo correctamente, confirmamos que el total de las cuatro frecuencias esperadas es igual al total de 250 estudiantes en la muestra:

$$85.104 + 111.896 + 22.896 + 30.104 = 250.000$$

También observe que los valores en el renglón “Total” y la columna “Total” son los mismos para las frecuencias observadas y las frecuencias esperadas por azar. Esto siempre debe ser el caso, proporcionando otra buena comprobación de su trabajo.

Tabla 10.5 Frecuencias observadas y frecuencias esperadas (entre paréntesis) para la tabla 10.3

	Biología	Administración	Total
Mujeres	32 (30.104)	110 (111.896)	142 (142.000)
Hombres	21 (22.896)	87 (85.104)	108 (108.000)
Total	53 (53.000)	197 (197.000)	250 (250.000)

Cálculo del estadístico ji cuadrada

Observe que las frecuencias esperadas en la tabla 10.5 parecen coincidir con las frecuencias observadas. Por ejemplo, la frecuencia esperada de alrededor de 85.1 para hombres con especialidad en administración es muy cercana a la frecuencia observada de 87. Por tanto, podríamos conjeturar que la información no nos da razón para rechazar la hipótesis nula de no relación entre las variables *género* y *especialidad*. Sin embargo, podemos ser más específicos en determinar una forma de cuantificar la diferencia entre las frecuencias observadas y las esperadas.

Denotemos las frecuencias observadas mediante *O* y las frecuencias esperadas mediante *E*. Esta notación, *O* – *E* (“*O* menos *E*”) indica la diferencia entre la frecuencia observada y la frecuencia esperada para cada celda. Estamos buscando una medida de la diferencia *total* para toda la tabla. No podemos dar tal medida simplemente sumando las diferencias individuales, *O* – *E*, ya que siempre suman cero. En lugar de eso, consideramos el *cuadrado* de la diferencia para cada celda (*O* – *E*)². Entonces hacemos cada valor de (*O* – *E*)² una diferencia relativa dividiéndolo entre la correspondiente frecuencia esperada; esto nos da la cantidad (*O* – *E*)²/*E* para cada celda. Sumando los valores individuales de (*O* – *E*)²/*E* nos da el **estadístico ji cuadrada**, denotado χ^2 (χ es la letra griega ji).

Determinación del estadístico ji cuadrada

Paso 1. Para cada celda en la tabla de dos entradas, identifique *O* como la frecuencia observada y *E* como la frecuencia esperada si la hipótesis nula es verdadera (no hay relación entre las variables).

Paso 2. Calcule el valor (*O* – *E*)²/*E* para cada celda.

Paso 3. Sume los valores del paso 2 para obtener el estadístico ji cuadrada.

$$\chi^2 = \text{suma de todos los valores } \frac{(O - E)^2}{E}$$

Entre mayor sea el valor de χ^2 , mayor será la diferencia promedio entre las frecuencias observadas y las esperadas en las celdas.

Para hacer estos cálculos de una manera organizada, es mejor construir una tabla como la 10.6, con un renglón para cada una de las celdas en la tabla de dos entradas original. Como se muestra en la celda inferior derecha para los datos género/especialidad el resultado es $\chi^2 = 0.350$.

Tabla 10.6 Cálculo del estadístico χ^2 para los datos de la tabla 10.5					
Resultado	O	E	O - E	(O - E) ²	(O - E) ² /E
Mujeres/administración	110	111.896	-1.896	3.595	0.032
Mujeres/biología	32	30.104	1.896	3.595	0.119
Hombres/administración	87	85.104	1.896	3.595	0.042
Hombres/biología	21	22.896	-1.896	3.595	0.157
Totales	250	250.000	0.000	14.380	$\chi^2 = 0.350$

UN MOMENTO DE REFLEXIÓN

¿Por qué los números en la columna O - E siempre deben sumar cero?

Cómo tomar la decisión

El valor de χ^2 nos da una manera de probar la hipótesis nula de no relación entre las variables. Si χ^2 es pequeña, entonces el promedio de diferencias entre las frecuencias observadas y las esperadas es pequeña, y *no* debemos rechazar la hipótesis nula. Si χ^2 es grande, entonces el promedio de diferencias entre las frecuencias observadas y las esperadas es grande y tenemos razón para rechazar la hipótesis nula de independencia. Para cuantificar el significado de “pequeña” y “grande”, comparamos el valor de χ^2 encontrado para la muestra con valores críticos:

- Si el valor calculado de χ^2 es *menor que* el valor crítico, las diferencias entre los valores observados y esperados son pequeñas y *no* hay evidencia suficiente para rechazar la hipótesis nula.
- Si el valor calculado de χ^2 es *mayor o igual al* valor crítico, entonces hay evidencia suficiente en la muestra para rechazar la hipótesis nula (al nivel de significancia dado).

La tabla 10.7 da los valores críticos de χ^2 para dos niveles de significancia, 0.05 y 0.01. Observe que los valores críticos difieren para distintos tamaños de tablas, así que debe estar seguro de que lee los valores críticos para un conjunto de datos en el renglón adecuado para el tamaño de la tabla. Para los datos género/especialidad que hemos estudiado en las tablas 10.3 y 10.5, hay dos renglones y dos columnas (no cuente los renglones o columnas de “total”), lo que significa una tabla de tamaño 2×2 . Al observar el primer renglón de la tabla 10.7, vemos que el valor crítico de χ^2 para significancia en el nivel 0.05 es 3.841. El valor de ji cuadrada que encontramos para los datos género/especialidad es $\chi^2 = 0.350$; puesto que es menor que el valor crítico de 3.841, no podemos rechazar la hipótesis nula. Por supuesto, no llegar a rechazar la hipótesis nula *no prueba* que especialidad y género sean independientes. Sólo significa que no tenemos evidencia suficiente para justificar el rechazo de la hipótesis nula de independencia.

Tabla 10.7 Valores críticos de χ^2 : rechazar H_0 sólo si $\chi^2 >$ valor crítico		
Tamaño de la tabla (renglones $\times$ columnas)	Nivel de significancia	
	0.05	0.01
2×2	3.841	6.635
2×3 o 3×2	5.991	9.210
3×3	9.488	13.277
2×4 o 4×2	7.815	11.345
2×5 o 5×2	9.488	13.277

NOTA TÉCNICA

El estadístico de prueba χ^2 es técnicamente una variable discreta, mientras que la distribución χ^2 es continua. La discrepancia no causa ningún problema sustancial mientras que la frecuencia esperada para cada celda sea por lo menos 5. Supondremos que esta condición se cumple para todos los ejemplos en este libro.

EJEMPLO 3 Prueba de vitamina C

Un estudio hipotético busca determinar si la vitamina C tiene un efecto en la prevención de resfriados. Entre una muestra de 220 personas, 105 personas elegidas aleatoriamente tomaron píldoras de vitamina C diariamente durante un periodo de 10 semanas, y las 115 personas restantes tomaron un placebo diariamente durante 10 semanas. Al finalizar las 10 semanas, el número de personas que habían tenido resfriados se registró. La tabla 10.8 resume los resultados. Determine si hay una relación entre tomar vitamina C y tener resfriado.

Tabla 10.8 Tabla de dos entradas para el número observado en cada categoría			
	Resfriado	No resfriado	Total
Vitamina C	45	60	105
Placebo	75	40	115
Total	120	100	220

Solución Iniciamos estableciendo las hipótesis nula y la alternativa.

- H_0 (hipótesis nula): no existe relación entre tomar vitamina C y atrapar un resfriado; esto es, la vitamina C no tiene más efecto sobre los resfriados que el placebo.
- H_a (hipótesis alternativa): existe una relación entre tomar vitamina C y atrapar un resfriado; esto es, el número de resfriados en los dos grupos no son lo que esperaríamos si la vitamina C y el placebo fueran igualmente eficaces (o igualmente ineficaces).

Como siempre, suponemos que la hipótesis nula es verdadera y calculamos la frecuencia esperada para cada celda en la tabla. Observando que el tamaño de la muestra es 220 y procediendo como en el ejemplo 2, encontramos las frecuencias esperadas siguientes:

Vitamina C y resfriado:

$$220 \times \underbrace{\frac{105}{220}}_{P(\text{vit. C})} \times \underbrace{\frac{120}{220}}_{P(\text{resfr.})} = 57.273$$

Vitamina C y no resfriado:

$$220 \times \underbrace{\frac{105}{220}}_{P(\text{vit. C})} \times \underbrace{\frac{100}{220}}_{P(\text{no resfr.})} = 47.727$$

Placebo y resfriado:

$$220 \times \underbrace{\frac{115}{220}}_{P(\text{placebo})} \times \underbrace{\frac{120}{220}}_{P(\text{resfr.})} = 62.727$$

Placebo y no resfriado:

$$220 \times \underbrace{\frac{115}{220}}_{P(\text{placebo})} \times \underbrace{\frac{100}{220}}_{P(\text{no resfr.})} = 52.273$$

La tabla 10.9 muestra la tabla de dos entradas con las frecuencias esperadas entre paréntesis.

Tabla 10.9 Frecuencias observadas y esperadas para el estudio de vitamina C			
	Resfriado	No resfriado	Total
Vitamina C	45 (57.273)	60 (47.727)	105 (105.000)
Placebo	75 (62.727)	40 (52.273)	115 (115.000)
Total	120 (120.000)	100 (100.000)	220 (220.000)

Ahora calculamos el estadístico ji cuadrada para los datos muestrales. La tabla 10.10 muestra cómo organizamos el trabajo; debe confirmar todos los cálculos mostrados.

Tabla 10.10 Tabla para calcular el estadístico χ^2 para el estudio de vitamina C

Resultado	O	E	O - E	(O - E) ²	(O - E) ² /E
Vitamina C/resfriado	45	57.273	-12.273	150.627	2.630
Vitamina C/no resfriado	60	47.727	12.273	150.627	3.156
Placebo/resfriado	75	62.727	12.273	150.627	2.401
Placebo/no resfriado	40	52.273	-12.273	150.627	2.882
Totales	220	220.000	0.000	602.508	$\chi^2 = 11.069$

Para tomar la decisión sobre rechazar la hipótesis nula, comparamos el valor de la ji cuadrada para los datos muestrales, $\chi^2 = 11.069$, con los valores críticos de la tabla 10.7. Revisamos el renglón para una tabla de tamaño 2×2 , puesto que los datos originales en la tabla 10.8 tenían dos renglones y dos columnas (sin contar los valores “total”). Vemos que el valor crítico de χ^2 para significancia al nivel 0.01 es 6.635. Puesto que nuestro valor muestral de $\chi^2 = 11.069$ es mayor que este valor crítico, rechazamos la hipótesis nula y concluimos que hay una relación entre la vitamina C y los resfriados. Esto es, con base en los datos de esta muestra existe razón para creer que la vitamina C *sí* tiene más efecto sobre los resfriados que un placebo.

A propósito...

Docenas de estudios cuidadosos se han realizado sobre la pregunta de la vitamina C y los resfriados. Algunos han encontrado altos niveles de confianza en los efectos de la vitamina C, pero otros no. A consecuencia de estos frecuentes resultados conflictivos, el tema de si la vitamina C ayuda a prevenir resfriados sigue siendo polémico.

EJEMPLO 4 Declarar o no declarar

La tabla de dos entradas 10.11 muestra cómo una declaración de culpable o no culpable afectó la sentencia en 1028 casos de robo elegidos de manera aleatoria en el área de San Francisco. Pruebe la afirmación de que la sentencia (prisión o no prisión) es independiente de la declaración.

Tabla 10.11 Frecuencias observadas para declaración y sentencia

	Prisión	No prisión	Total
Declararse culpable	392	564	956
Declararse no culpable	58	14	72
Total	450	578	1028

Fuente: *Law and Society Review*, Vol. 16, No. 1.

Solución Las hipótesis nula y alternativa para el problema son:

H_0 : (hipótesis nula): la sentencia en casos de robo es independiente de la declaración.

H_a : (hipótesis alternativa): la sentencia en casos de robo depende de la declaración.

Encontramos el número de personas *esperado* siguiente en cada categoría, suponiendo que las variables de los renglones son independientes de las variables de las columnas:

$$\text{Culpable y prisión: } 1028 \times \frac{956}{1028} \times \frac{450}{1028} = 418.482$$

$$\text{Culpable y no prisión: } 1028 \times \frac{956}{1028} \times \frac{578}{1028} = 537.518$$

$$\text{No culpable y prisión: } 1028 \times \frac{72}{1028} \times \frac{450}{1028} = 31.518$$

$$\text{No culpable y no prisión: } 1028 \times \frac{72}{1028} \times \frac{578}{1028} = 40.482$$

La tabla 10.12 resume las frecuencias observadas y las frecuencias esperadas.

Tabla 10.12 Frecuencias observadas y frecuencias esperadas por azar (en paréntesis) para declaración y sentencia

	Prisión	No prisión	Total
Culpable	392 (418.482)	564 (537.518)	956
No culpable	58 (31.518)	14 (40.482)	72
Total	450	578	1028

Como es usual, calculamos χ^2 organizando nuestro trabajo como se muestra en la tabla 10.13, donde la *O* denota frecuencia observada y *E* denota la frecuencia esperada. Como se muestra en la celda inferior derecha, el estadístico ji cuadrada para estos datos es $\chi^2 = 42.556$. Este valor es mucho mayor que los valores críticos (para una tabla de 2×2) para significancia en el nivel 0.05 ($\chi^2 = 3.841$) y el nivel 0.01 ($\chi^2 = 6.635$). Por tanto, los datos respaldan rechazar la hipótesis nula y aceptar la hipótesis alternativa. Con base en estos datos, hay razón para creer que la sentencia dada en un caso de robo está asociada con la declaración. En específico, las personas que se declararon culpables, en realidad menos fueron a prisión que lo esperado (por azar) y más evitaron la prisión que lo esperado. De las personas que se declararon no culpables, en realidad más fueron a prisión de lo esperado y menos evitaron la prisión de lo esperado. Recuerde que la prueba no demuestra una relación causal entre la declaración y la sentencia.

Tabla 10.13 Cálculo de χ^2

Resultado	<i>O</i>	<i>E</i>	<i>O</i> - <i>E</i>	(<i>O</i> - <i>E</i>) ²	(<i>O</i> - <i>E</i>) ² / <i>E</i>
Culpable/prisión	392	418.482	-26.482	701.296	1.676
Culpable/no prisión	564	537.518	26.482	701.296	1.305
No culpable/prisión	58	31.518	26.482	701.296	22.251
No culpable/no prisión	14	40.482	-26.482	701.296	17.324
Totales	1028	1028.000	0.000	2805.184	$\chi^2 = 42.556$

UN MOMENTO DE REFLEXIÓN

Si usted fuese el abogado para un sospechoso de robo, ¿cómo podría el resultado del ejemplo anterior afectar su estrategia en la defensa de su cliente? Explique.

Sección 10.2 Ejercicios

Alfabetización estadística y pensamiento crítico

- Tablas de dos entradas.** ¿Qué es una tabla de dos entradas y qué son los valores que se ingresan en una tabla de dos entradas?
- Relación entre variables.** Suponga que hemos analizado las frecuencias en una tabla de dos entradas con género (hombre/mujer) correspondiendo a los renglones y mano dominante (izquierda/derecha) correspondiendo a las columnas. También suponga que rechazamos la independencia entre género y mano dominante. ¿Podemos concluir que el género de una persona tiene un efecto si la persona es zurda o derecha? ¿Por qué sí o por qué no?

- Frecuencia esperada.** Suponga que conocemos las frecuencias en una tabla de dos entradas con género (hombre/mujer) como la variable del renglón y multa por exceso de velocidad (sí/no) como la otra variable. Dado que las frecuencias observadas deben ser números enteros, ¿las frecuencias esperadas también deben ser números enteros?
- Notación.** En el contexto de tablas de dos entradas, describa las notaciones siguientes: *O*, *E* y χ^2 .

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. Encuesta. En una encuesta rápida se les pregunta a los sujetos si pueden identificar el año en el cual Estados Unidos declaró su independencia. Los resultados de esa pregunta se resumen en una tabla de dos entradas.

6. Estadístico de prueba χ^2 . Una tabla de dos entradas se utiliza para calcular un estadístico de prueba y se obtuvo el valor $\chi^2 = -2.500$.

7. Estadístico de prueba χ^2 . En una tabla de dos entradas, todas las frecuencias observadas son muy cercanas a las frecuencias esperadas, por lo que el estadístico de prueba χ^2 es muy pequeño y no llega a rechazar la hipótesis nula de independencia entre las variables del renglón y de la columna.

8. Hipótesis nula. Suponga que una tabla de dos entradas está configurada para que *género* (hombre/mujer) represente la variable del renglón y *respuesta* (sí/no) a una pregunta de la encuesta, represente la variable de la columna. La hipótesis nula es el enunciado de que el género y la respuesta son independientes.

Conceptos y aplicaciones

Resultados de la encuesta. En los ejercicios 9 al 12 suponga que a una muestra aleatoria simple, de hombres y mujeres, se presenta una encuesta de preguntas con respuesta sí/no y los resultados se resumen en una tabla de dos entradas con el formato de la tabla siguiente. Utilice el valor dado del estadístico de prueba χ^2 y el nivel de significancia para probar la independencia entre género y respuesta.

	Sí	No
Mujer		
Hombre		

9. Estadístico de prueba: $\chi^2 = 0.051$; nivel de significancia: 0.01.

10. Estadístico de prueba: $\chi^2 = 10.785$; nivel de significancia: 0.01.

11. Estadístico de prueba: $\chi^2 = 2.924$; nivel de significancia: 0.05.

12. Estadístico de prueba: $\chi^2 = 15.238$; nivel de significancia: 0.05.

Prueba de hipótesis completa. En los ejercicios 13 al 20 realice los pasos siguientes:

- Indique la hipótesis nula y la alternativa.
- Suponiendo independencia entre las dos variables, determine la frecuencia esperada para cada celda de la tabla.
- Encuentre el valor del estadístico de prueba χ^2 .
- Utilice el nivel de significancia dado para encontrar el valor crítico de χ^2 .
- Usando el nivel de significancia dado, complete la prueba para la afirmación de que las dos variables son independientes. Indique la conclusión que aborda la afirmación original.

13. Demografía estudiantil. La tabla siguiente (consistente en datos nacionales) muestra la distribución de estudiantes de tiempo parcial y de tiempo completo para una muestra de 148 hombres y mujeres. Utilice un nivel de significancia de 0.05 para probar la afirmación de independencia entre género y categoría de estudiante (tiempo completo o tiempo parcial).

	Tiempo completo	Tiempo parcial
Mujeres	47	37
Hombres	38	26

14. Concurrencia de votantes. La tabla siguiente muestra el número de ciudadanos en una muestra que votaron en la pasada elección presidencial, de acuerdo al género (consistente con los datos de la población nacional). Utilice un nivel de significancia de 0.05 para probar la afirmación de que el género es independiente de la asistencia de votantes.

	Votaron	No votaron
Mujeres	140	120
Hombres	130	110

15. Correo electrónico y privacidad. A los trabajadores y jefes de nivel superior se les preguntó si era muy poco ético monitorear el correo electrónico de los empleados; los resultados se resumen en la tabla siguiente (con base en datos de una encuesta de Gallup). Utilice un nivel de significancia de 0.05 para probar la afirmación de que la respuesta es independiente de si el sujeto es trabajador o jefe de nivel superior.

	Sí	No
Trabajadores	192	244
Jefes	40	81

16. Eficacia de cascos para bicicleta. Un estudio realizado entre 531 personas lesionadas en accidentes de bicicleta, muestra elegida aleatoriamente, se resume en la tabla siguiente (con base en datos de “A Case-Control Study of the Effectiveness of Bicycle Safety Helmets in Preventing Facial Injury” de Thompson, Thompson, Rivara y Wolf *American Journal of Public Health*, volumen 80, número 12). Con un nivel de significancia de 0.01, pruebe la afirmación de que el uso de un casco es independiente de recibir lesiones faciales.

	Se usó casco	Sin casco
Lesiones faciales recibidas	30	182
No hay lesiones faciales	83	236

17. Tratamiento de artritis. A 98 participantes se les dio un nuevo tratamiento experimental para la artritis, 56 mostraron mejoría. De los 92 participantes que se les dio un placebo, 49 mostraron mejoría. Construya una tabla de dos entradas para estos datos, y luego utilice un nivel de significancia de 0.05 para probar la afirmación de que la

mejoría es independiente de que al participante se le haya dado el medicamento o un placebo.

18. Bebida y embarazo. Una muestra aleatoria simple de 1 252 mujeres embarazadas menores de 25 años incluye 13 que estaban bebiendo alcohol durante su embarazo. Una muestra aleatoria simple de 2 029 mujeres embarazadas de 25 años y más de edad incluye a 37 que estaban bebiendo alcohol durante su embarazo. (Los datos tienen como base resultados del Centro Nacional para Estadísticas de la Salud de Estados Unidos). Utilice un nivel de significancia de 0.05 para probar la afirmación de que la categoría de edad (menor de 25 y de 25 o mayores) es independiente de que la mujer embarazada estuviera bebiendo durante el embarazo.

19. Crimen y desconocidos. La tabla siguiente lista los resultados de encuestas obtenidas de una muestra aleatoria de diferentes víctimas de crímenes (con base en datos del Departamento de Justicia de Estados Unidos). Utilice un nivel de significancia de 0.01 para probar la afirmación de que el tipo de crimen es independiente de si el criminal era un extraño.

	Homicidio	Robo	Asalto
El criminal era un desconocido	12	379	727
El criminal era un conocido o un pariente	39	106	642

20. Fumar en China. La tabla siguiente resume los resultados de una encuesta de hombres de 15 años o mayores que viven en el distrito Minhang de China (con base en datos de “Cigarette Smoking in China” por Gong, Koplan, Feng, *et al.* *Journal of the American Medical Association*, volumen 274, número 15). Los hombres se clasifican por su curso de estudios actual y si fuman. Con un nivel de significancia de 0.05, pruebe la afirmación de que fumar es independiente del nivel educativo.

	Escuela primaria	Escuela secundaria	Nivel superior
Fumador	606	1234	100
Nunca ha fumado	205	505	137



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 10 en www.aw.com/bbt.

21. Construcción de tablas de dos entradas. Seleccione dos variables que al parecer tengan una relación que valga la pena investigar. Una variable debe tener al menos dos categorías de individuos, por ejemplo, dos o más categorías de edades, categorías raciales o ubicaciones geográficas. La otra variable debe tener al menos dos categorías para algún factor social, económico o de salud, por ejemplo, dos o más categorías de ingresos, categorías de bebida, o categorías de logros educativos. Encuentre la población requerida o datos muestrales para llenar una tabla de dos en-

tradas para las dos variables. Analice si parece existir una relación entre las variables. Un buen lugar para iniciar es el sitio web del *Statistical Abstract of the United States* de la oficina de Censo de Estados Unidos.

22. Análisis de tablas de dos entradas. Seleccione dos variables que parezcan tener una relación que valga la pena investigar. Una variable debe tener al menos dos categorías de individuos, por ejemplo, dos o más categorías de edades, categorías raciales o ubicaciones geográficas. La otra variable debe tener al menos dos categorías para algún factor social, económico o de salud, por ejemplo, dos o más categorías de ingresos, categorías de bebida, o categorías de logros educativos. Encuentre las frecuencias de datos necesarias para llenar una tabla de dos entradas para las dos variables. Realice una prueba de hipótesis para determinar si existe una relación entre las variables. Una buena fuente es el sitio web del *Statistical Abstract of the United States* de la oficina de Censo de Estados Unidos.



EN LAS NOTICIAS



23. Tablas de dos entradas en las noticias. Es raro (pero no imposible) ver una tabla de dos entradas en un artículo periodístico. Pero con frecuencia una nota periodística proporciona información que podría expresarse en una tabla de dos entradas. Encuentre un artículo que analice una relación entre dos variables que podría expresarse en una tabla de dos entradas. Cree la tabla.

24. Prueba de hipótesis en las noticias. Con frecuencia las noticias describen el resultado de estudios estadísticos en que las conclusiones provienen de una prueba de hipótesis que involucra tablas de dos entradas. Sin embargo, los reportes raramente proporcionan la tabla real o describen los detalles de la prueba de hipótesis. Encuentre una noticia reciente en la que usted crea que las conclusiones *probablemente* estuvieron basadas en una prueba de hipótesis con una tabla de dos entradas. Suponiendo que esté en lo correcto, describa en palabras cómo, quizá, se llevó a cabo la prueba de hipótesis. Esto es, describa la hipótesis nula y la alternativa y el procedimiento mediante el cual los investigadores, quizá, llevaron a cabo su prueba.

25. Su propia prueba de hipótesis. Piense en un ejemplo de algo que le gustaría conocer que podría probarse con una prueba de hipótesis en una tabla de dos entradas. Sin recolectar realmente los datos ni realizar cálculo alguno, describa cómo podría abordar su estudio. Esto es, describa cómo podría recolectar los datos, explique cómo los organizaría en una tabla de dos entradas, indique la hipótesis nula y la alternativa que aplicaría y cómo realizaría la prueba de hipótesis para llegar a una conclusión.

10.3 Análisis de varianza (ANOVA de un factor)

Hasta ahora hemos examinado pruebas de hipótesis con tres tipos diferentes de afirmaciones para la hipótesis nula: afirmaciones de que una media poblacional es igual a algún valor ($H_0: \mu = \text{valor afirmado}$), las cuales examinamos con una distribución normal en la sección 9.2 y con la distribución t en la sección 10.1; afirmaciones de que una proporción poblacional es igual a algún valor ($H_0: p = \text{valor afirmado}$), que examinamos en la sección 9.3, y afirmaciones de que dos variables son independientes entre sí (H_0 : no hay relación), que analizamos en la sección 10.2. Los estadísticos han desarrollado técnicas para considerar muchos otros tipos de hipótesis nulas, haciendo posible aplicar la estadística a un increíble rango de aplicaciones. Para darle una probada de lo que es posible con estadística —y quizá para animarlo a estudiar más estadística— consideramos brevemente un tipo más de prueba de hipótesis en esta última sección del libro.

Prueba de hipótesis para varianza

Una muestra aleatoria simple de 12 páginas se obtuvo de cada uno de tres libros diferentes: *El oso y el dragón* de Clancy, *Harry Potter y la piedra filosofal* de J. K. Rowling, y *La guerra y la paz* de León Tolstoi. La puntuación de facilidad de lectura de Flesch se obtuvo para cada una de esas páginas y los resultados se listan en la tabla 10.14. El sistema de puntuación de facilidad de lectura de Flesch da como resultado puntuaciones *más altas* para el texto que es *más sencillo* de leer. Puntuaciones bajas están asociadas con trabajos difíciles de leer. En esta sección nuestro objetivo es utilizar estos datos muestrales de sólo 12 páginas de cada libro para hacer inferencias acerca de la legibilidad de la *población* de todas las páginas de cada libro.

En una exploración informal de los datos muestrales podemos investigar el centro, la variación, la distribución y los valores extremos. La tabla 10.15 muestra los estadísticos muestrales importantes para nuestro caso. Si usted compara los datos originales en la tabla 10.14 y las medias muestrales en la tabla 10.15, notará que aunque unas cuantas puntuaciones están más lejos de la media que la mayoría de las demás (tal como la puntuación más baja de Clancy de 43.9 y la puntuación más baja de Rowling de 70.9), ninguno de los valores parecen ser tan extremos para considerarlos valores extremos. Además, el estudio detallado de los datos sugiere que las muestras provienen de poblaciones con distribuciones que son cercanas a la normal.

Tabla 10.14 Puntuaciones de facilidad de lectura de Flesch

Clancy	Rowling	Tolstoi
58.2	85.3	69.4
73.4	84.3	64.2
73.1	79.5	71.4
64.4	82.5	71.6
72.7	80.2	68.5
89.2	84.6	51.9
43.9	79.2	72.2
76.3	70.9	74.4
76.4	78.6	52.8
78.9	86.2	58.4
69.4	74.0	65.4
72.9	83.7	73.6

Tabla 10.15 Estadísticas para puntuaciones de legibilidad

	Puntuación de facilidad de lectura de Flesch		
	Clancy	Rowling	Tolstoi
Tamaño de la muestra n	12	12	12
Media muestral $\bar{x}$	70.73	80.75	66.15
Desviación estándar muestral s	11.33	4.68	7.86

Incluso antes de estudiar los datos podríamos esperar que el libro de Rowling sea el más sencillo de leer de los tres libros, ya que es el único escrito para niños. Del mismo modo, podríamos esperar que el libro de Tolstoi sea el más difícil de leer ya que es una traducción de un clásico ruso. Ahora, revise las puntuaciones medias de legibilidad en la tabla 10.15. Recordando que una puntuación más alta de Flesch indica un nivel más sencillo de lectura, los datos parecen respaldar nuestras expectativas. Rowling tiene la puntuación más alta de legibilidad y Tolstoi la más baja. Incluso, las tres medias muestrales no son tan diferentes, variando sólo de 66.15 para Tolstoi a 80.75 para Rowling, y para cada caso el tamaño muestral es relativamente pequeño $n = 12$. Por tanto, llegamos a nuestra pregunta estadística clave para esta sección: ¿estos datos muestrales nos proporcionan evidencia suficiente para concluir que los libros de Clancy, Rowling y Tolstoi en realidad tienen puntuaciones medias de Flesch diferentes?

Para responder esta pregunta seguimos los mismos principios generales presentados para pruebas de hipótesis en la sección 9.1. Para empezar, identificamos la hipótesis nula. Puesto que queremos saber si los tres libros en realidad tienen diferentes puntuaciones medias de Flesch, iniciamos con la suposición de que éstas *no* tienen medias diferentes. En otras palabras, nuestra hipótesis nula es que las puntuaciones medias de Flesch para los tres libros son iguales. Entonces, la hipótesis alternativa es que las tres medias son diferentes. La prueba de hipótesis debe decirnos si se rechaza o no se rechaza la hipótesis nula. Rechazar la hipótesis nula nos permitiría concluir que los libros en realidad tienen puntuaciones medias de Flesch diferentes, como esperamos. No rechazar la hipótesis nula nos diría que los datos no proporcionan evidencia suficiente para concluir que las puntuaciones medias de Flesch son diferentes.

Recuerde que para este ejemplo cada media poblacional (μ) representa la puntuación media de Flesch que obtendríamos si midiésemos la puntuación para *todas* las páginas en cada libro. Por tanto, usando notación más formal podemos escribir la hipótesis nula como

$$H_0: \mu_{\text{Clancy}} = \mu_{\text{Rowling}} = \mu_{\text{Tolstoi}}$$

Como puede ver, necesitamos una prueba de hipótesis que nos permitirá determinar si tres poblaciones diferentes tienen la misma media. El método que usamos se denomina **análisis de varianza**, comúnmente abreviado **ANOVA** (por sus siglas en inglés). El nombre proviene del estadístico formal conocido como *varianza* de un conjunto de valores muestrales, como señalamos brevemente en la sección 4.3, la varianza se define como el cuadrado de la desviación estándar muestral, o s^2 . (Por ejemplo, si una muestra de estaturas tiene una desviación estándar $s = 3.0$ cm, su varianza es $s^2 = 9.0$ cm²).

NOTA TÉCNICA

El método de análisis de varianza puede usarse con dos medias, pero es equivalente a una prueba t que junta las dos varianzas muestrales. Existe otra prueba t que puede usarse con dos muestras independientes, y por lo regular lo hace mejor. Ninguna de estas pruebas t están incluidas en este libro.

Definición

El **análisis de varianza (ANOVA)** es un método para probar la igualdad de tres o más medias poblacionales mediante el análisis de las varianzas muestrales.

En específico, el método usado para analizar datos como los de la tabla 10.14 se denomina análisis de varianza de *una entrada* (ANOVA de un factor), ya que los datos muestrales están separados en grupos de acuerdo a *una* sola característica (o factor). En este ejemplo, la característica es el autor (Clancy, Rowling o Tolstoi). También hay un método conocido como *análisis de varianza de dos factores* que permite las comparaciones entre poblaciones separadas en categorías por dos características (o factores). Por ejemplo, podríamos separar las estaturas de las personas usando las dos características siguientes: (1) género (masculino o femenino) y (2) derecho o zurdo. En este libro no consideramos análisis de varianza de dos factores.

Cómo realizar la prueba

El análisis de varianza está basado en este concepto fundamental: *supongamos* que las poblaciones tienen la misma varianza y entonces comparamos la varianza *entre* las muestras con la varianza *dentro* de las muestras. En específico, el estadístico de prueba (usualmente llamado F) para análisis de varianza de un factor es la razón de esas dos varianzas:

$$\text{estadístico de prueba } F \text{ (para ANOVA de un factor)} = \frac{\text{varianza entre las muestras}}{\text{varianza dentro de las muestras}}$$

Los cálculos reales de este estadístico de prueba son tediosos, por lo que en la actualidad casi siempre son hechos con programas estadísticos (vea la sección Uso de la tecnología, página 437). Sin embargo, podemos interpretar el estadístico como sigue, usando nuestro ejemplo de la legibilidad de los tres libros:

- La varianza *entre* las muestras es una medida de cuánto difieren las medias muestrales (de la tabla 10.15) una de otra.
- La varianza *dentro* de las muestras es una medida de cuánto difieren una de otra las puntuaciones de facilidad de lectura de Flesch para las 12 páginas en cada muestra individual (de la tabla 10.14).

- Si las tres medias poblacionales en realidad fuesen todas iguales —como lo afirma la hipótesis nula— entonces esperaríamos que la media muestral de cualquier muestra individual caiga dentro del rango de variación de cualquier otra muestra individual. El estadístico de prueba ($F = \text{varianza entre muestras} / \text{varianza dentro de muestras}$) nos dice si es el caso.

Un estadístico de prueba grande nos dice que las medias muestrales difieren *más* que los datos dentro de las muestras individuales, que sería *poco probable* si las medias poblacionales en realidad fueran iguales (como lo afirma la hipótesis nula). Esto es, un estadístico de prueba grande proporciona evidencia para rechazar la hipótesis nula que las medias poblacionales son iguales.

Un estadístico de prueba pequeño nos dice que las medias muestrales difieren *menos* que los datos dentro de las muestras individuales, sugiriendo que la diferencia entre las medias muestrales con facilidad podrían haber surgido por azar. Por tanto, un estadístico de prueba pequeño no proporciona evidencia suficiente para rechazar la hipótesis nula de que las medias poblacionales son iguales.

Observe que el estadístico de prueba F para análisis de varianza desempeña un papel semejante al de la puntuación estándar z o el estadístico t en pruebas de hipótesis consideradas antes. Por tanto, al igual que lo hicimos en los casos previos, cuantificamos la interpretación del estadístico de prueba encontrando su valor P , que nos proporciona la probabilidad de obtener resultados muestrales al menos tan extremos como aquéllos obtenidos suponiendo que la hipótesis nula sea verdadera (todas las medias poblacionales son iguales). Un valor P pequeño muestra que es poco probable que obtuviésemos los resultados muestrales por azar con medias poblacionales iguales. Un valor P grande muestra que con facilidad podríamos obtener los resultados muestrales por azar con medias poblacionales iguales.

Al igual que el estadístico de prueba mismo, el cálculo del valor P por lo general se hace con ayuda de software. Una vez que hemos encontrado el valor P , lo utilizamos para tomar la decisión final, igual que lo hacemos en otros casos: si el valor P es pequeño (tal como 0.05 o menor), rechazamos la hipótesis nula de medias iguales; si el valor P es grande (por ejemplo, mayor que 0.05), no rechazamos la hipótesis nula de medias iguales. El cuadro siguiente resume los requisitos para el análisis de varianza de un factor y el procedimiento apoyado en software esbozado en esta sección.

ANOVA de un factor para probar $H_0: \mu_1 = \mu_2 = \mu_3 = \dots$

Paso 1. Ingrese los datos muestrales en un paquete de software estadístico y utilice el software para determinar el estadístico de prueba ($F = \text{varianza entre muestras} / \text{varianza dentro de muestras}$) y el valor P del estadístico de prueba.

Paso 2. Tome la decisión de rechazar o no rechazar la hipótesis nula con base en el valor P del estadístico de prueba:

- Si el valor P es menor o igual al nivel de significancia, rechazar la hipótesis nula de medias iguales y concluya que al menos una de las medias es diferente de las otras.
- Si el valor P es mayor que el nivel de significancia, no rechazar la hipótesis nula de medias iguales.

El método es válido mientras se cumplan los requisitos siguientes. Las poblaciones tienen distribuciones que son aproximadamente normales con la misma varianza, y las muestras de cada población son muestras aleatorias simples que son independientes una de la otra.

EJEMPLO 1 Legibilidad de Clancy, Rowling, Tolstoi

Dadas las puntuaciones de legibilidad listados en la tabla 10.14 y un nivel de significancia de 0.05, pruebe la hipótesis nula de que las tres muestras provienen de poblaciones con medias que son iguales.

Solución Iniciamos verificando los requisitos para usar el análisis de varianza de un factor. Como observamos anteriormente, una revisión cuidadosa de los datos sugiere que cada muestra proviene de una distribución que es aproximadamente normal. Las desviaciones estándar

muestrales no son muy diferentes, por lo que es razonable suponer que las tres poblaciones tienen la misma varianza. Las muestras son aleatorias simples y todas ellas son independientes. Por tanto los requisitos se satisfacen.

Ahora probamos la hipótesis nula: las medias poblacionales son iguales ($H_0: \mu_1 = \mu_2 = \mu_3$). La sección Uso de la tecnología en la página 437 describe cómo calcular el estadístico de prueba y el valor *P* con varios paquetes de cómputo. La tabla siguiente muestra la exhibición resultante con Excel; otros paquetes de cómputo proporcionarán exhibiciones semejantes.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	<i>F</i>	Probabilidad	Valor crítico para <i>F</i>
Entre grupos	1338.002222	2	669.0011111	9.469487401	0.000562133	3.284924333
Dentro de grupos	2331.386667	33	70.64808081			
Total	3669.388889	35				

Observe que la pantalla incluye las columnas para *F* y para el valor *P* (rotulada como Probabilidad). Aquí, estos dos elementos son de interés para nosotros, los cuales interpretamos a continuación:

- *F* es el estadístico de prueba para el análisis de varianza de un factor (*F* = varianza entre muestras/varianza dentro de las muestras). Note que es mucho mayor que 1, lo cual indica que las medias muestrales diferirían más de lo que esperaríamos si todas las medias poblacionales fuesen iguales.
- El valor *P* nos dice la probabilidad de haber obtenido tal resultado extremo como resultado del azar, si la hipótesis nula es verdadera. Observe que el valor *P* es extremadamente pequeño, mucho menor que el valor de 0.05 necesario para rechazar la hipótesis nula al nivel 0.05 de significancia (y también mucho menor que el 0.01 necesario para rechazar al nivel de significancia 0.01).

Concluimos que existe evidencia suficiente para rechazar la hipótesis nula, esto es, los datos muestrales respaldan la afirmación de que las tres medias poblacionales no son todas iguales. Con base en páginas elegidas aleatoriamente de *El oso y el dragón* de Clancy, *Harry Potter y la piedra filosofal*, de Rowling y *La guerra y la paz* de Tolstoi, concluimos que esos libros tienen niveles de legibilidad que no son iguales. Observe que *no* hemos concluido que estos libros tienen los niveles de legibilidad que esperamos, Rowling como la más sencilla y Tolstoi el más difícil, puesto que la hipótesis muestra sólo que las legibilidades no son iguales. Sin embargo, nuestra expectativa parece razonable ya que las medias muestrales en la tabla 10.15 van en el orden esperado.

Sección 10.3 Ejercicios

Alfabetización estadística y pensamiento crítico

1. **ANOVA.** ¿Qué método representa ANOVA, y ¿cuál es el propósito del método?
2. **Varianza.** ¿Qué es la varianza de una muestra?
3. **Comparación de especialidades.** Una estudiante en el Colegio de Newport aplica un examen de razonamiento abstracto para seleccionar aleatoriamente especialidades de inglés, matemáticas y ciencias en su colegio. Luego utiliza un análisis de varianza para concluir que las puntuaciones medias no son todas iguales. ¿Ella puede concluir que, en Estados Unidos, las especialidades de inglés, matemáticas y ciencias tienen puntuaciones medias de razonamiento abstracto que no todas son iguales? ¿Por qué sí o por qué no?
4. **ANOVA de un factor.** ¿Por qué el método de esta sección es conocido como análisis de varianza de *un factor*?

¿Tiene sentido? Para los ejercicios 5 al 8 decida si el enunciado tiene sentido (o es claramente verdadero) o no tiene sentido (o claramente es falso). Explique claramente; no todos los enunciados tienen respuestas definitivas, por lo que su explicación es más importante que la respuesta elegida.

5. **Valor *P*.** En una prueba para la igualdad de los niveles medios de colesterol de personas de Estados Unidos, Noruega y China, los datos muestrales dieron como resultado un valor *P* de 0.99, por lo que rechazamos la igualdad de las tres medias poblacionales.
6. **Valor *P*.** En una prueba para igualdad de las edades medias de miembros de la audiencia de una película animada, una comedia y una de suspenso, el valor *P* es 0.009, por lo que rechazamos la igualdad de las tres medias poblacionales.

7. **ANOVA.** Para un estudio de puntuaciones del CI de niños se seleccionaron familias con tres hijos. Una muestra incluye puntuaciones del CI del hijo que nació primero, la segunda muestra incluye las puntuaciones del CI del hijo que nació en segundo lugar y la tercera muestra incluye las puntuaciones del CI del hijo que nació en tercer lugar. El método de análisis de varianza puede usarse para probar la igualdad de las tres medias poblacionales.
8. **ANOVA.** Un fabricante de defensas para automóviles está probando tres máquinas diferentes y usa análisis de varianza para comparar las cantidades de variación resultantes de las tres máquinas diferentes.

Conceptos y aplicaciones

9. **Legibilidad de autores.** El ejemplo en esta sección utilizó las puntuaciones de facilidad de lectura de Flesch para páginas elegidas aleatoriamente de libros de Tom Clancy, J. K. Rowling y León Tolstoi. Pero cuando se utilizan las puntuaciones de nivel de calidad de Flesch-Kincaid, los resultados de análisis de varianza de STATDISK son los que se muestran en la figura 10.2. Suponga que queremos usar un nivel de significancia de 0.05 al probar la hipótesis nula de que los tres autores tienen puntuaciones de nivel de calidad de Flesch-Kincaid con las mismas medias:
- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la hipótesis alternativa?
 - Identificar el valor P .
 - Con base en los resultados precedentes, ¿qué concluiría acerca de la igualdad de las medias poblacionales?

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	2	68.187222	34.093611	8.978506	3.284914	0.00077
Error:	33	125.309167	3.797247			
Total:	35	193.496389	5.528468			

Reject the Null Hypothesis
Reject equality of means

Figura 10.2

10. **Pruebas de inflamabilidad en tejidos realizadas en laboratorios diferentes.** La prueba semicontrolada vertical fue usada para realizar pruebas de inflamabilidad en prendas de dormir infantiles. Las piezas de tela fueron quemadas bajo condiciones controladas. Después que se paró de quemar, se midió y registró la longitud de la parte carbonizada. Las mismas muestras de tejidos fueron probadas en cinco laboratorios diferentes. Los resultados del análisis de varianza de Excel 2003 se muestran a continuación.
- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la hipótesis alternativa?
 - Identifique el valor P .
 - ¿Hay evidencia suficiente para respaldar la afirmación de que las medias para los diferentes laboratorios no son todas iguales? Suponga que se usa un nivel de significancia 0.05.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	2.087194264	4	0.521798566	2.949333035	0.030665893	2.588834036
Dentro de grupos	7.607597403	43	0.17692087			
Total	9.694791667	47				

- 11. Tiempos de maratón.** Una muestra aleatoria de hombres que terminaron el maratón de Nueva York se divide en tres categorías con edades de 21-29, 30-39 y 40 o mayores. Los tiempos (en segundos) se obtuvieron para los hombres seleccionados. Los resultados obtenidos del análisis de varianza de Excel 2003 se muestran a continuación.

- ¿Cuál es la hipótesis nula?
- ¿Cuál es la hipótesis alternativa?
- Identifique el valor *P*.
- ¿Existe evidencia suficiente para respaldar la afirmación de que los hombres de diferentes categorías de edad tienen diferentes tiempos medios?

<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>Grados de libertad</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
Entre grupos	3532063.284	2	1766031.642	0.188679406	0.828324293	3.080387501
Dentro de grupos	1010875649	108	9359959.71			
Total	1014407712	110				

- 12. Presión sanguínea sistólica en grupos diferentes de edad.** Una muestra aleatoria de 40 mujeres se divide en tres categorías diferentes con edades de 20, 20 a 40 y más de 40. Los resultados obtenidos del análisis de varianza con SPSS se muestran a continuación.

- ¿Cuál es la hipótesis nula?
- ¿Cuál es la hipótesis alternativa?
- Identifique el valor *P*.
- ¿Hay evidencia suficiente para respaldar la afirmación de que las mujeres en las diferentes categorías de edad tienen distintos niveles medios de presión sanguínea?

	<i>Suma de cuadrados</i>	<i>g.l.</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Sig.</i>
Entre grupos	937.930	2	468.965	1.655	.205
Dentro de los grupos	10484.470	37	283.364		
Total	11422.400	39			

En los ejercicios 13 al 16 utilice software para realizar la prueba de análisis de varianza.

- 13. Lesión en la cabeza en un accidente automovilístico.** En experimentos de accidentes automovilísticos realizados por la Administración Nacional de Seguridad en el Transporte, automóviles nuevos fueron comprados y estrellados contra una barrera fija a 35 millas por hora. Los automóviles subcompactos fueron Ford Escort, Honda Civic, Hyundai Accent, Nissan Sentra y Saturn SL4. Los automóviles compactos fueron Chevrolet Cavalier, Dodge Neon, Mazda 626 DX, Pontiac Sunfire y Subaru Legacy. Los automóviles medianos fueron Chevrolet Camaro, Dodge Intrepid, Ford Mustang, Honda Accord y Volvo S70. Los automóviles grandes fueron Audi A8, Cadillac Deville, Ford Crown Victoria, Oldsmobile Aurora y Pontiac Bonneville. La información de lesión en la cabeza para los maniqués en el asiento del conductor se lista a continuación. Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las diferentes categorías de peso tienen la misma media. ¿Los datos muestrales sugieren que automóviles más grandes son más seguros?

<i>Subcompacto:</i>	681	428	917	898	420
<i>Compacto:</i>	643	655	442	514	525
<i>Mediano:</i>	469	727	525	454	259
<i>Grande:</i>	384	656	602	687	360

- 14. Desaceleración en el pecho en un accidente automovilístico.** La información de desaceleración en el pecho (en g, aceleración debida a la gravedad) de las pruebas descritas en el ejercicio 13 se da a continuación. Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las diferentes categorías de peso tienen la misma media. ¿Los datos sugieren que los automóviles grandes son más seguros?

Subcompacto: 55 47 59 49 42
Compacto: 57 57 46 54 51
Mediano: 45 53 49 51 46
Grande: 44 45 39 58 41

- 15. Arqueología: ancho del cráneo de diferentes épocas.** Los valores en la tabla siguiente son las medidas de anchos máximos (en milímetros) de cráneos egipcios de diferentes épocas (con base en datos de *Ancient Races of the Thebaid*, de Thomson y Randall-Maciver). Cambios en la forma de la cabeza al paso del tiempo sugieren que la mezcla de razas ocurrió con las poblaciones inmigrantes. Utilice un nivel de significancia de 0.05 para probar la afirmación de que las diferentes épocas tienen la misma media.

4000 a. de C.	1850 a. de C.	150 d. de C.
131	129	128
138	134	138
125	136	136
129	137	139
132	137	141
135	129	142
132	136	137
134	138	145
138	134	137

- 16. Energía solar en diferentes climas.** Una estudiante de uno de los autores vive en una casa con un sistema eléctrico solar. A la misma hora cada día recolecta las lecturas de voltaje de un medidor conectado al sistema; los resultados se listan en la tabla siguiente. Utilice un nivel de significancia de 0.05 para probar la afirmación de que la media de las lecturas es la misma para los tres tipos diferentes de días. Podríamos esperar que un sistema solar proporcione más energía eléctrica en días soleados que en días nublados

o días lluviosos. ¿Hay evidencia suficiente para respaldar una afirmación de medias poblacionales diferentes?

Días soleados	Días nublados	Días lluviosos
13.5	12.7	12.1
13.0	12.5	12.2
13.2	12.6	12.3
13.9	12.7	11.9
13.8	13.0	11.6
14.0	13.0	12.2



Proyectos para internet y más allá

Para enlaces útiles seleccione “Links for Internet Projects” para el capítulo 10 en www.aw.com/bbt.

- 17. Personalidades notables.** El libro *World Almanac and Book of Facts* incluye una sección denominada “Personalidades notables” con subsecciones para arquitectos, artistas, líderes de negocios, humoristas y otras categorías. Diseñe y realice un estudio que empiece con la selección de muestras de grupos elegidos, seguida de la comparación de medias de longevidades de las personas de categorías diferentes. ¿Algún grupo parece tener longevidad que sea diferente de los otros grupos?
- 18. ANOVA.** Encuentre un artículo de una revista que se refiera al uso de análisis de varianza. Identifique la prueba que fue usada y describa la conclusión. ¿El resultado de la prueba tiene como resultado rechazar la igualdad de medias? ¿Cuál fue el valor P ? ¿Cuál fue el papel que desempeñó el método de análisis de varianza?



EN LAS NOTICIAS



- 19. Deportes.** Encuentre una noticia reciente en la que se discuta sobre los salarios en diferentes equipos profesionales, tal como equipos de béisbol. Encuentre los salarios y utilice análisis de varianza para probar la hipótesis nula de medias iguales. Resuma sus hallazgos y escriba un reporte que incluya sus conclusiones.

Ejercicios de repaso del capítulo

1. Un estudio patrocinado por AT&T y la Asociación Estadounidense de Automóviles incluyó los datos muestrales de la tabla siguiente:

	Tuvo accidente el año pasado	No tuvo accidente el año pasado
Usuario de teléfono celular	23	282
No es usuario de teléfono celular	46	407

- Compare el porcentaje de usuarios de teléfono celular que tuvieron accidente con el porcentaje de aquellos que no son usuarios de teléfono celular y no tuvieron accidente. Con base en estos resultados, ¿parece que los teléfonos celulares son peligrosos?
- Identifique la hipótesis nula y la hipótesis alternativa para probar la afirmación de que tener un accidente es independiente del uso de teléfono celular.
- Encuentre el valor esperado para cada celda de la tabla, suponiendo que tener un accidente es independiente del uso de teléfono celular.
- Encuentre el valor del estadístico χ^2 para una prueba de hipótesis de la afirmación de que tener un accidente es independiente del uso de teléfono celular.
- Con base en los resultados del inciso d y el tamaño de la tabla, consulte la tabla 10.7 (en la página 423) y determine qué se sabe acerca del valor *P*.
- Con base en los resultados anteriores, ¿qué puede concluir de la prueba de hipótesis acerca de si las dos variables (*uso de teléfono celular* y *tener un accidente*) son independientes?
- ¿La conclusión del inciso f es consistente con lo que ahora se sabe acerca del uso de teléfono celular y manejo?

2. La carga axial de una lata es el peso máximo que soporta lateralmente, y debe ser mayor que 165 libras porque es la fuerza máxima aplicada cuando la tapa superior es presionada en casa. A continuación se listan las cargas axiales (en libras) para una muestra aleatoria de latas de aluminio de 12 onzas.

270 273 258 204 254 228 282

Utilice un nivel de significancia de 0.05 para probar la afirmación de que la muestra es de una población con una media mayor que 165 libras.

- Usando la muestra de datos del ejercicio 2, construya un intervalo de confianza al 95% para estimar la media de carga axial de todas las latas.
- En la tabla siguiente están listadas las temperaturas corporales (°F) de sujetos elegidos aleatoriamente de tres grupos diferentes de edades. La pantalla de STATDISK en la figura 10.3 resulta de estos valores muestrales. Suponga que queremos usar un nivel de significancia de 0.05 para probar la afirmación de que los tres grupos de edades tienen la misma temperatura media corporal.

18-20	21-29	30 y mayores
98.0	99.6	98.6
98.4	98.2	98.6
97.7	99.0	97.0
98.5	98.2	97.5
97.1	97.9	97.3

- ¿Cuál es la hipótesis nula?
- ¿Cuál es la hipótesis alternativa?
- Identifique el valor *P*.
- ¿Existe evidencia suficiente para rechazar la afirmación de que los tres grupos de edades tienen la misma temperatura media corporal?

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	2	1.729333	0.864667	1.87971	3.88529	0.194915
Error:	12	5.52	0.46			
Total:	14	7.249333	0.51781			

Figura 10.3

Cuestionario del capítulo

- Se obtuvo una muestra aleatoria simple de 15 valores de una población distribuida normalmente con una desviación estándar desconocida. ¿Cuál de las distribuciones siguientes es la más apropiada para una prueba de hipótesis que incluya una afirmación acerca de una media poblacional?
 - distribución normal
 - distribución t
 - distribución ji cuadrada
 - distribución uniforme
- Se obtuvo una muestra aleatoria simple de 45 valores de una población distribuida normalmente con una desviación estándar desconocida. ¿Cuál de las distribuciones siguientes es la más apropiada para una prueba de hipótesis que incluya una afirmación acerca de una media poblacional?
 - distribución normal
 - distribución t
 - distribución ji cuadrada
 - distribución uniforme
- Se obtuvo una muestra aleatoria simple de 15 valores de una población distribuida normalmente con una desviación estándar conocida. ¿Cuál de las distribuciones siguientes es la más apropiada para una prueba de hipótesis que incluya una afirmación acerca de una media poblacional?
 - distribución normal
 - distribución t
 - distribución ji cuadrada
 - distribución uniforme
- ¿Cuál es la hipótesis nula para una afirmación acerca de que la puntuación media en un examen del SAT es mayor que 1500?
- ¿Cuál es la hipótesis alternativa para una afirmación acerca de que la puntuación media en un examen del SAT es mayor que 1500?
- Determine si el enunciado siguiente es verdadero o falso: una prueba t se utiliza para probar la afirmación de que, en una tabla de dos entradas, la variable del renglón y la variable de la columna tienen alguna relación.
- Suponga que quiere probar la afirmación de que los estudiantes con especialidad en ciencias, literatura y administración tienen la misma puntuación media del CI. ¿Qué método usaría para probar esa afirmación?
- Si la prueba de hipótesis de la afirmación descrita en el ejercicio 7 da como resultado un valor P de 0.3500, ¿qué concluiría de la hipótesis nula?
- Una tabla de dos entradas, fundada en los resultados de una encuesta, consiste en dos reglones que representan sexo (masculino/femenino) y dos columnas que representan la respuesta a una pregunta (sí/no). ¿Cuál es la hipótesis nula para una prueba que determine si existe alguna relación entre sexo y respuesta?
- Si la prueba de hipótesis descrita en el ejercicio 9 da como resultado un valor P de 0.001, ¿qué concluiría acerca de la hipótesis nula?

Uso de la tecnología

Intervalos de confianza usando la distribución t

SPSS: ingrese los datos muestrales en la ventana del editor de datos de SPSS. Dé clic en **Analyze**, luego seleccione **Descriptive Statistics**, seleccione el elemento del menú **Explore**. Dé clic en la variable de la izquierda del cuadro de diálogo y dé clic en el botón de enmedio para pegarla en el cuadro "Dependent List". En el mismo cuadro de diálogo dé clic en el botón **Statistics** y asegúrese que el cuadro "Descriptives" esté marcado y el nivel de confianza tenga el valor deseado. Dé clic en **Continue**. Dé clic en **OK**. Se mostrarán las cotas inferior y superior del intervalo de confianza en el cuadro rotulado con "Descriptives".

Excel: utilice el complemento Data Desk XL que es un complemento para Excel.

Primero ingrese los datos muestrales en la columna A. Si está usando Excel 2003, dé clic en **DDXL**. Si está usando Excel 2007, dé clic en **Add-Ins**, luego en **DDXL**. Seleccione **Confidence Intervals**. Debajo de las opciones de tipos de función, seleccione **1 Var t Intervalo**, si σ no es conocida. (Si σ es conocida, seleccione **1 Var z Intervalo**). Dé clic en el icono del lápiz e ingrese el rango de datos, tal como A1:A12, si tiene 12 valores listados en la columna

A. Dé clic en **OK**. En el cuadro de diálogo seleccione el nivel de confianza. (Si está usando **1 var z Intervalo**, también ingrese el valor de σ). Dé clic en **Compute Interval** y aparecerá el intervalo de confianza.

STATDISK: seleccione **Analysis**, luego **Confidence Intervals**, luego **Mean - One Sample**. En el cuadro de diálogo que aparece, primero ingrese el nivel de significancia como un número decimal. Ingrese 0.95 para un nivel de confianza del 95%. Proceda a ingresar los otros elementos que se piden y luego dé clic en **Evaluate**; aparecerá el intervalo de confianza.

Pruebas de hipótesis con la distribución t

SPSS: ingrese los datos muestrales en la ventana del editor de datos de SPSS. Dé clic en **Analyze**, luego seleccione el elemento del menú **Compare Means**, y luego seleccione **One-Sample T Test**. Dé clic en la variable en la izquierda del cuadro de diálogo y dé clic en el botón de enmedio para pegarla en el cuadro "Test Variable(s)". Ingrese el valor supuesto de la media poblacional en el cuadro "Test Value". (Éste es el mismo valor usado en la hipótesis nula). Dé clic en **OK**. La pantalla incluirá los componentes clave en el cuadro rotulado "One-Sample Test". Ese cuadro incluirá el valor del estadístico de prueba t en la columna

con el encabezado “*t*”. Ese mismo cuadro incluirá un valor *P* en la columna rotulada con “Sig. (2-tailed)”. Observe que esto es el valor correcto del valor *P* para una prueba de dos colas, pero debe ajustarse para una prueba de una cola. Para pruebas de una cola por lo regular puede dividir el valor mostrado entre 2, pero esto no funciona si la hipótesis alternativa no es consistente con la media muestral, como es en cualquiera de estos dos casos: (1) una prueba de cola izquierda con una media muestral es mayor que la media supuesta en H_0 , o bien (2) una prueba de cola derecha con una media muestral menor que la media supuesta en H_0 .

Excel: Excel no tiene una función integrada para una prueba *t*, por lo que utilice el complemento Data Desk XL que es un complemento para Excel.

Primero ingrese los datos muestrales en la columna A. Si está utilizando Excel 2003, dé clic en **DDXL**. Si está usando Excel 2007, dé clic en **Add-Ins**, luego dé clic en **DDXL**. Seleccione **Hypothesis Tests**. Debajo de las opciones del tipo de función, seleccione **1 Var *t* Test**. Dé clic en el icono del lápiz e ingrese el rango de valores de datos, tal como A1:A12, si tiene 12 valores listados en la columna A. Dé clic en **OK**. Siga los cuatro pasos listados en el cuadro de diálogo. Después de dar clic en **Compute** en el paso 4, obtendrá el valor *P*, estadístico de prueba y la conclusión.

DATADISK: seleccione **Analysis**, luego seleccione **Hypothesis Testing**. Para los métodos analizados en la sección 10.1 seleccione **Mean - One Sample**. Aparecerá un cuadro de diálogo. Dé clic en el cuadro en la esquina superior izquierda y seleccione el elemento que coincide con la afirmación de que se probará. Proceda a ingresar los otros elementos en el cuadro de diálogo, y luego dé clic en **Evaluate**. Los resultados incluirán el estadístico de prueba y el valor *P*.

Pruebas de hipótesis con tablas de dos entradas

SPSS: en la ventana de datos de SPSS ingrese todos los conteos de frecuencias en la primera columna, ingrese los nombres de la columna correspondiente en la segunda columna, e ingrese los nombres de los renglones en la tercera columna. Los conteos de frecuencias deben ponderarse como sigue: dé clic en **Data**, luego seleccione el elemento del menú **Weight Cases**. Dé clic en el botón rotulado con “Weight cases by” y pegue la variable que representa las frecuencias en la caja rotulada con “Frequency Variable”. Ahora dé clic en **Analyze**, seleccione **Descriptive Statistics** y seleccione **Crosstabs**. Dé clic en la variable que representa los nombres de los renglones y péguela en el cuadro rotulado con “Row(s)”. Dé clic en la variable que representa los nombres de las columnas y péguela en el cuadro rotulado con “Column(s)”. Ahora dé clic en el botón **Statistics** y proceda a verificar el cuadro rotulado con “Chi-Square”. Luego dé clic en **Continue**. Dé clic en el botón **Cells** y proceda a verificar las cajas etiquetadas “Observed” y “Expected”. Dé clic en el botón **Continue**. Por último, dé clic en **OK**. Una de las cajas de diálogo mostrará las frecuencias observadas y las esperadas. La caja rotulada con “Pearson Chi Square” y el valor *P* están en el mismo renglón.

Excel: utilice el complemento DDXL.

Primero ingrese los *nombres* de las categorías de las columnas en la columna A. Segundo, ingrese los *nombres* de las categorías en la columna B. Tercero, ingrese las frecuencias observadas correspondientes en la columna C. Si está usando Excel 2003 dé clic en **DDXL**. Si está usando Excel 2007 dé clic en **Adds-Ins**. Luego dé clic en **DDXL**. Ahora seleccione **Tables**, luego **Indep. Test for Summ Data**. Dé clic en el icono del lápiz para **Variable One Names** e ingrese el rango de celdas que contienen los nombres de las categorías de *columnas*, tal como A1:A6. Dé clic en el icono del lápiz para **Variable Row Names** e ingrese el rango de celdas que contienen los nombres de las categorías de los renglones, tal como B1:B6. Dé clic en el icono del lápiz para **Counts** e ingrese el rango de celdas que contienen las frecuencias observadas, tal como C1:C6. Dé clic en **OK** para obtener los resultados de la prueba, que incluyen el estadístico de prueba χ^2 y el valor *P*. (Las frecuencias esperadas también se mostrarán).

STATDISK: primero ingrese las frecuencias observadas en las columnas de la ventana de datos. Seleccione **Analysis** de la barra del menú principal, luego seleccione **Contingency Tables**, y proceda a identificar las columnas que contienen las frecuencias. Dé clic en **Evaluate**. Los resultados de STATDISK incluyen el estadístico de prueba, el valor crítico *P* y la conclusión.

Análisis de varianza

SPSS: ingrese todos los datos en la primera columna en la ventana de datos de SPSS. En la segunda columna ingrese los correspondientes *números* para las categorías. Ingrese 1 para los valores de la primera categoría, ingrese 2 para los valores de la segunda categoría y así sucesivamente. Dé clic en **Analyze**, seleccione **Compare Means**, luego seleccione **One-Way ANOVA**. En el cuadro de diálogo que aparece, dé clic en la variable para la primera columna de datos y dé clic en el botón de enmedio para moverla al cuadro rotulado como “Dependent List”. También dé clic en la variable que contiene los números de categoría y dé clic en el botón de enmedio para moverla al cuadro rotulado “Factor”. Dé clic en **OK** para obtener los resultados, que incluirán el valor *P*.

Excel: primero ingrese los datos en las columnas A, B, C,... Si está usando Excel 2003 dé clic en **Tools** (Herramientas) en la barra del menú principal, luego seleccione **Data Analysis** (Análisis de datos). Si está usando Excel 2007, dé clic en **Data** (Datos), luego dé clic en **Data Analysis** (Análisis de datos). Seleccione **Anova: Single Factor**. En el cuadro de diálogo ingrese el rango que contiene los datos muestrales. (Por ejemplo, ingrese A1:C12, si el primer valor está en el renglón 1 de la columna A y la última entrada está en el renglón 12 de la columna C). Los resultados incluirán el valor *P* para la prueba de análisis de varianza.

STATDISK: ingrese los datos en las columnas de la ventana de datos. Seleccione **Analysis** de la barra de menú principal, luego seleccione **One-Way Analysis of Variance** y proceda a seleccionar las columnas de los datos muestrales. Cuando lo haya hecho, dé clic en **Evaluate**. Los resultados incluirán el valor *P* para la prueba de análisis de varianza.

HABLEMOS DE CRIMINOLOGÍA

¿Puede descubrir un fraude cuando lo ve?

Suponga que su profesor le da una tarea en la que usted debe lanzar una moneda 200 veces y registrar los resultados en orden. Los dos conjuntos de datos siguientes representan resultados entregados por dos alumnos. Ahora suponga que sabe que uno de los estudiantes en realidad hizo la tarea, mientras que el otro falsificó los datos. ¿Puede decir cuál es falso?

Conjunto de datos 1 (H = cara; T = cruz)

H T H T H H T T T T T H T H T T T H
H H T T T H T T H T T H H H T T H H T
H T H T H H H H T T H T H T H H H H T H
T T H H T T H H H T T T T T H H H T H T
T H T H T T H H T T H T H T H H T T H T
T T H T H T H H T T T H H T T H H H H T
H T H H T H T T T T H T T T H T H T H H
T H T H H H H T H T H H H T T T H T T H
T H T T T H T H H T H H H H T T T H H T
T H T H H T T H T H H H T H H T T H

Conjunto de datos 2 (H = cara; T = Cruz)

T H H T T H H T T H T H H T T H H T H T
H T T H T T H H T T H T T T H H T H T H
H H T T H T H H T H T T H T T H H T T H
H T H H T T H T H T H H T H T H T H H T
H T H T T H T T H H T T H H T H T H H T
T H T H H T T H T H T T H T T H T H H T
H T H T T H T H H T H T H T H T H H T H
T T H T H T H H T T H T T H T H H T H H
H T H H T T H T H T H T H T H H T H T T
T H H T H T H H T H H T T H H T H T H T

Para hacer el trabajo un poco más sencillo, la tabla siguiente resume características de los dos conjuntos de datos que podrían ayudarle a decidir cual es falsa.

Características del conjunto de datos 1	Características del conjunto de datos 2
Total de 97 H, 103 T Dos casos de 6 T seguidas Cinco casos de 4 H seguidas Tres casos de 4 T seguidas	Total de 101 H, 99 T Ningún caso de más de 3 H o 3 T seguidas



Si usted es como la mayoría de la gente, quizá conjeture que el conjunto de datos 1 es el falso. Después de todo, el número total de caras y cruces están muy lejos de las 100 de cada una que muchas personas esperan, y tiene dos casos en los que hubo 6 cruces seguidas, además de varios casos en los que hubo 4 caras o cruces seguidas.

Pero considere esto: la probabilidad de obtener 6 caras seguidas es $(1/2)^6$ o 1 en 64. La posibilidad de 6 cruces consecutivas también es 1 en 64. Por tanto, con 200 lanzamientos, la posibilidad de obtener al menos un caso de 6 caras seguidas es muy buena, por lo que cadenas de caras y cruces consecutivas en el conjunto de datos 1 en realidad no son sorprendentes. En contraste, el conjunto de datos no tiene cadenas más largas que 4 de caras o cruces seguidas, aunque la probabilidad de tal cadena es sólo $(1/2)^4$, o 1 en 16, concluimos que casi seguramente el conjunto de datos 2 es falso.

Este ejemplo simple revela una aplicación importante de la estadística en criminología: con frecuencia es posible atrapar a gente que ha falsificado datos de cualquier clase. Por ejemplo, la estadística ayuda a los reguladores de bancos y auditores financieros a atrapar estados financieros fraudulentos, el Servicio de Recaudación Interno identifica gente con declaraciones de impuestos falsa, y científicos para atrapar a otros científicos han falsificado datos.

Una de las herramientas más poderosas para detectar fraudes fue identificada por el físico Frank Benford. En la década de los treinta del siglo pasado, Benford observó que las tablas de logaritmos (que los científicos e ingenieros usaban regularmente en los días anteriores a las calculadoras) tendían a ser más usadas en las páginas iniciales, donde los números iniciaban con el dígito 1, que en las páginas finales. Continuando con esta observación, pronto descubrió que muchos conjuntos de números de la vida diaria, tal como valores del mercado de valores, estadísticas de béisbol y áreas de lagos, incluyen más números iniciales con el dígito 1 que con el dígito 2, y más iniciando con 2 que con 3 y así sucesivamente. Eventualmente, él publicó una fórmula para describir con qué frecuencia inician los números con diferentes dígitos, y esta fórmula ahora se denomina *ley de Benford*. La figura 10.4 muestra lo que esta ley predice para los primeros dígitos, junto con resultados reales de varios conjuntos reales de números. Observe cuán bien la ley de Benford describe los resultados. (Una curiosidad, la ley de Benford fue descubierta más de 50 años antes y publicada por el astrónomo y matemático Simon Newcomb en 1881. Sin embargo, el artículo de Newcomb había sido olvidado en la época que Benford hizo su trabajo).

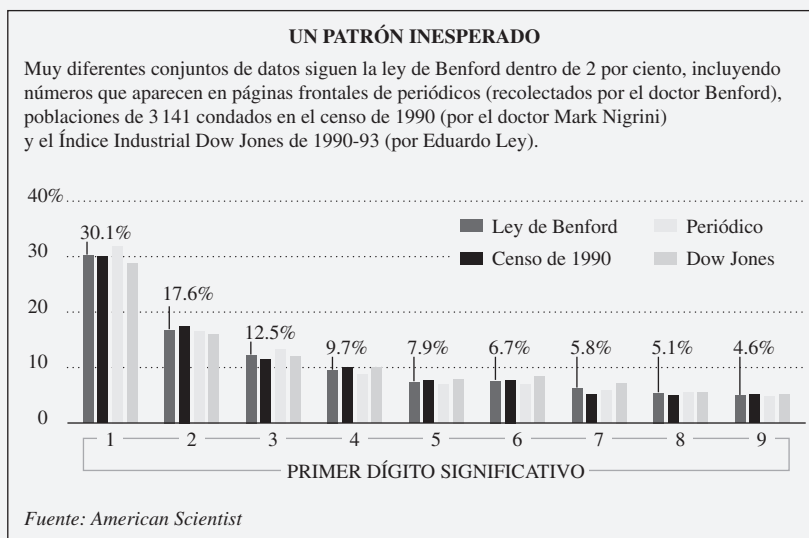


Figura 10.4 Fuente: Malcolm W. Browne, "Following Benford's Law, or Looking Out for No. 1", *New York Times*, 4 de agosto de 1998, página B10.

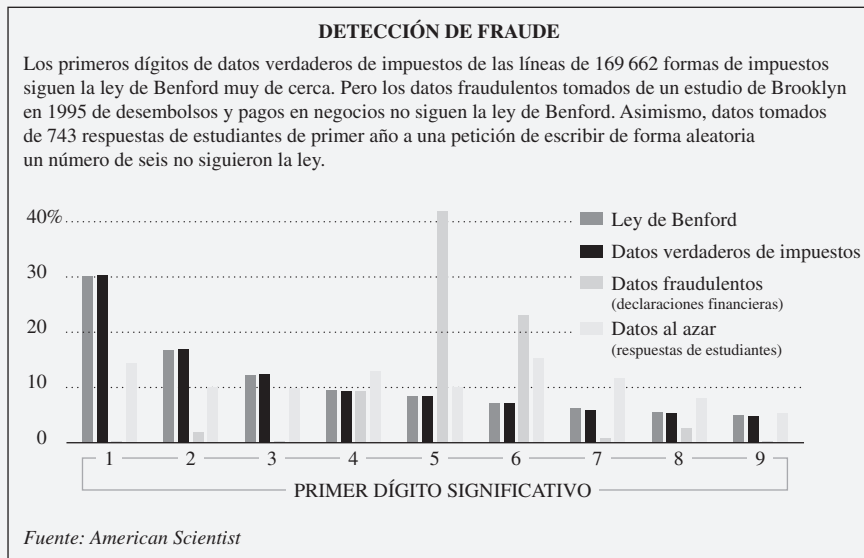


Figura 10.5 Fuente: Malcolm W. Browne, “Following Benford’s Law, or Looking Out for No. 1”, *New York Times*, 4 de agosto de 1998, página B10.

La ley de Benford es sorprendente, ya que la mayoría de la gente supone que cada dígito (1 a 9) sería igualmente probable como dígito inicial. En realidad es el caso para conjuntos de números aleatorios tales como números de lotería, que no es más probable que inicien con 1 que con cualquier otro dígito. (Por lo que la ley de Benford no debe usarse para seleccionar números de lotería). Sin embargo, puesto que la ley de Benford se aplica a muchos conjuntos de *datos reales*, puede usarse para detectar fraude. Como se muestra en la figura 10.5, los datos reales de formas de impuestos (en barras negras) siguen muy de cerca la ley de Benford (barras gris fuerte). En contraste, los datos de un conjunto de declaraciones financieras (barras gris) examinadas en un estudio en 1995 no siguen la ley de Benford. Con base en este hecho, el fiscal del Distrito sospechó de fraude, que eventualmente fue capaz de demostrar. Los “datos aleatorios” (barras gris claro) provienen de datos de estudiantes del profesor Theodore P. Hill en el Instituto Tecnológico de Georgia (la fuente de mucha de la información para esta sección). Las conjeturas no siguen la ley de Benford, razón por la cual la gente que trata de falsificar datos con frecuencia es atrapada.

La ley de Benford desconcertó a científicos y matemáticos por décadas. Ahora parece que está bien entendida, aunque aún es difícil de explicar. A continuación está una explicación de por qué la ley de Benford se aplica al Índice Industrial Dow Jones, gracias al doctor Mark J. Nigrini de la Universidad Metodista del Sur (como se publicó en el *New York Times*): imagine que el Dow estará en 1 000, por lo que el primer dígito es 1 y sube a una tasa de aproximadamente 20% anual. El tiempo de duplicación a esta tasa de aumento es un poco menos de 4 años, por lo que el Dow permanecerá en los mil, aún con el primer dígito 1, durante casi cuatro años, hasta que llegue a 2 000. Entonces tendría un primer dígito 2 hasta que llegue a 3 000. Sin embargo, moverse de 2 000 a 3 000 requiere de un aumento de sólo 50%, lo cual tarda sólo un poco más de dos años. Así, el primer dígito 2 ocurrirá por sólo un poco más de la mitad de días que un primer dígito 1 ocurrió. Cambios subsecuentes tardan aún menos tiempo. En el momento en que el Dow Jones llegue a 9 000 toma sólo 11% de aumento y sólo siete meses para llegar a la marca de 10 000, por lo que un primer dígito de 9 sólo ocurre durante siete meses. Sin embargo, en la marca de 10 000, el Dow Jones regresa otra vez a un primer dígito de 1, y no cambiaría hasta que duplique el Dow Jones a 20 000, lo cual significa otros casi cuatro años a la tasa de aumento de 20% anual. Por tanto, si grafica el número de días que el Dow Jones tiene cada dígito inicial del 1 al 9, encontrará que el número 1 sería el dígito inicial por periodos más largos que el número 2, y así disminuyendo en sucesión.

En resumen, la ley de Benford muestra que los números no siempre surgen con la frecuencia que la mayoría de la gente supondría. En consecuencia, no sólo ayuda a explicar muchos misterios acerca de los números (tal como los del Dow Jones), sino también se ha convertido en una herramienta valiosa para la detección de fraudes criminales.

PREGUNTAS PARA DISCUSIÓN

1. Intente este experimento con sus amigos: pídale a un amigo que registre 200 lanzamientos reales de una moneda y a otro amigo que intente falsificar datos de 200 lanzamientos de una moneda. Que le entreguen a usted sus resultados de manera anónima, de modo que usted no sepa cuál hoja proviene de cuál amigo. Con las ideas de esta sección intente determinar cuál conjunto de datos es real y cuál es falso. Después que haga su conjetura verifique con sus amigos para ver si estuvo en lo correcto. Analice su capacidad para detectar fraudes.
2. Analice brevemente cómo la ley de Benford podría usarse para detectar declaraciones de impuestos fraudulentas.
3. ¿La sola ley de Benford puede probar qué datos son fraudulentos? O, ¿sólo puede señalar datos que deben ser investigados con mayor detalle? Explique.
4. Encuentre un conjunto de datos similares a los conjuntos mostrados en la figura 10.4 y construya un diagrama de barras que muestre las frecuencias de los dígitos 1 a 9. ¿Los datos siguen la ley de Benford?

LECTURAS SUGERIDAS

Browne, Malcolm W., "Following Benford's Law, or Looking Out for No. 1", *New York Times*, 4 de agosto de 1998.

Hill, Theodore P., "The First Digit Phenomenon", *American Scientist*, volumen 86, número 358, julio-agosto de 1998.

Nigrini, Mark K., *Digital Analysis Using Benford's Law*, Global Audit Publications, Vancouver, 2000.

¿Qué puede hacer una alumna de cuarto grado con estadística?

Emily Rosa de nueve años de edad estaba en cuarto grado, tratando de decidir qué hacer para su proyecto escolar de ciencias. Ella pensaba en hacer un proyecto sobre los colores de los chocolates M&M, cuando observó que su madre, una enfermera, veía una videocinta acerca de una práctica denominada “toque terapéutico” o “TT”. El TT es un destacado tratamiento médico alternativo, practicado en muchos lugares en todo el mundo e incluso enseñado en algunas escuelas de enfermería. Pero ninguna prueba estadística válida ha mostrado claramente si en realidad funciona. Emily le dijo a su mamá que tenía una idea para probar el TT y quería hacerlo para su trabajo de ciencias.

A pesar del nombre los terapeutas del TT en realidad *no* tocan a sus pacientes, ellos mueven sus manos a unas cuantas pulgadas por arriba del cuerpo del paciente. Los que apoyan el toque terapéutico afirman que estos movimientos con las manos permiten a los terapeutas entrenar para sentir y manipular, lo que ellos llaman un “campo de energía humano”. Al hacer estas manipulaciones de manera apropiada, supuestamente los terapeutas pueden curar muchos achaques y enfermedades. El proyecto para la feria de la ciencia de Emily Rosa buscaba determinar si terapeutas entrenados en TT en realidad podrían sentir un campo de energía humano.

Para hacer su proyecto, Emily reclutó a 21 terapeutas de TT para participar en un experimento sencillo. Cada terapeuta, en una mesa, sentado enfrente de Emily con sus brazos descansando en la mesa y con las palmas hacia arriba. Entonces, Emily coloca una división de cartón con huecos para los brazos del terapeuta. Esto evita que Emily y el terapeuta se vean cara a cara, pero permitió a Emily ver las manos del terapeuta.

Entonces, Emily coloca una de sus manos a unas cuantas pulgadas por arriba de una de las dos manos del terapeuta. Si el terapeuta en verdad puede sentir el campo de “energía humano” de Emily, entonces él sería capaz de decir si su mano derecha o izquierda estaba más cerca de la mano de Emily. Por tanto, cada ensayo del experimento terminaba al registrar si el terapeuta estaba en lo correcto o no en la identificación de la mano.

Emily tomó varias precauciones para garantizar que su experimento sería estadísticamente válido. Por ejemplo, para asegurar que su elección entre las dos manos fuese aleatoria, Emily utilizó el resultado del lanzamiento de una moneda para determinar en cada caso si colocaba su mano sobre la mano izquierda o la mano derecha del terapeuta. Y para asegurarse que tenía suficiente información para evaluar la significancia estadística, 14 de los 21 terapeutas hicieron 10 ensayos cada uno, mientras que 7 hicieron 20 ensayos cada uno.

Los resultados fueron una falla penosa para los terapeutas TT. Ya que sólo hubo dos posibles en cada ensayo —mano izquierda o mano derecha— por puro azar los terapeutas tenían que ser capaces de adivinar la mano correcta en alrededor de 50% de las veces. Pero los resultados generales mostraron que obtuvieron la respuesta correcta sólo 44% de las veces. Además, ningún terapeuta tuvo un desempeño estadísticamente significativo mejor al esperado por el azar. Emily también comprobó para ver si los terapeutas con más experiencia lo hacían mejor que los menos experimentados. No fue así. La conclusión de Emily: si existe una cosa como “campo de energía humano” (lo cual ella duda), los terapeutas TT no pueden sentirlo. Y aún si existe el “campo de energía humano”, es difícil imaginar cómo los terapeutas TT podrían usarlo para curar, si ellos no pueden detectar su presencia.

Uno de los aspectos más interesantes de este estudio fue que Emily fue capaz de hacerlo en su totalidad. Otros escépticos del TT habían querido realizar estudios similares en el pasado, pero los terapeutas TT habían rechazado participar. Un escéptico famoso, el mago James



Randi, había ofrecido un premio de \$1.1 millones a cualquier terapeuta TT que pudiese pasar un examen semejante al de Emily. Sólo una persona aceptó el reto de Randi, sólo tuvo éxito en 11 de 20 ensayos, casi lo mismo que se esperaría por azar. Así, ¿por qué Emily fue capaz de tener éxito donde investigadores experimentados habían fallado? Al parecer, los terapeutas accedieron a participar en el experimento de Emily porque no se sentían amenazados por una niña de cuarto grado.

La novedad en el proyecto para la feria de la ciencia de Emily atrajo la atención de los medios, y no pasó mucho tiempo antes de que se enterase el psiquiatra retirado de Pennsylvania Stephen Barrett. El doctor Barrett especializado en echar por tierra terapias “milagrosas”, y él convencieron a Emily y a su mamá para reportar sus resultados en un artículo de investigación médica. El artículo fue publicado en el *Journal of the American Medical Association* (1 de abril de 1998) cuando Emily tenía 11 años, convirtiéndola en la autora más joven de un artículo en esa prestigiosa revista.

PREGUNTAS PARA DISCUSIÓN

1. Después que los resultados de Emily fueron publicados, muchos partidarios del TT aseguraron que su experimento era inválido ya que ella y su madre estaban sesgadas en contra del TT. Con base en la manera en que su experimento fue diseñado, ¿considera que su sesgo personal podría haber afectado sus resultados? ¿Por qué sí o por qué no?
2. Otra objeción al experimento de Emily fue que sólo era ciego sencillo en lugar de doble ciego. Esto es, la terapeuta no podría ver lo que Emily estaba haciendo, pero Emily podía ver lo que la terapeuta estaba haciendo. ¿Usted cree que, en este caso, esta objeción es válida? ¿Podría pensar en una manera en que el experimento de Emily podría repetirse pero haciéndolo doble ciego?
3. El experimento de Emily no fue una prueba directa de si el tratamiento TT funcionaba, ya que no comprobó si los pacientes en realidad mejoraban cuando eran tratados mediante TT. Sugiera una manera estadísticamente válida para probar si el TT es más eficaz que un placebo.
4. Con base en los resultados de Emily, ahora los escépticos dicen que TT es tan claramente no válido que ya no debe ser utilizado ni financiado. ¿Está de acuerdo? ¿Por qué sí o por qué no?

LECTURAS SUGERIDAS

Ball, T. S. y Alexander, D. D., “Catching Up with Eighteenth Century Science in the Evaluation of Therapeutic Touch”, *Skeptical Inquirer*, julio/agosto de 1998.

Kolata, Gina, “4th Grader Challenges Alternative Therapy”, *New York Times*, 2 de abril de 1998.

Rosa, E., “TT and Me”, *Skeptic Magazine*, septiembre de 1998.

Rosa, L., Rosa, E., Sarner, L. y Barret, S., “A Close Look at Therapeutic Touch”, *Journal of the American Medical Association*, volumen 279, número 1005, 1 de abril de 1998.

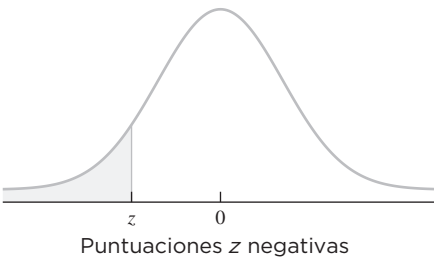
Epílogo: una perspectiva sobre estadística

Un solo curso introductorio de estadística no puede transformarlo en un estadístico experto. Después de estudiar estadística en este libro puede sentir que aún no ha dominado el material al nivel necesario para usar la estadística en aplicaciones reales. Sin embargo, por ahora, debe entender lo suficiente sobre estadística para interpretar de manera crítica los reportes de investigación estadística que ve en las noticias y para conversar con expertos en estadística cuando necesite más información. Y, si va a tomar un curso adicional de estadística, debe estar bien preparado para entender temas importantes que están fuera del alcance de este libro introductorio.

Lo que es más importante, aunque este libro no está diseñado para convertirlo en un experto en estadística sí lo está para hacerlo una persona mejor educada y con una mejor posibilidad de trabajo en comercio. Debe entender y conocer los conceptos básicos de probabilidad y aleatoriedad. Debe saber que en el camino de obtener conocimiento sobre un conjunto de datos es importante investigar medidas de tendencia central (como la media y la mediana), medidas de variación (como rango y desviación estándar), la naturaleza de la distribución (vía una tabla de frecuencias o una gráfica) y la presencia de datos extremos. Debe saber y entender la importancia de la estimación de parámetros poblacionales (como una media poblacional o una proporción poblacional), así como pruebas de hipótesis acerca de parámetros poblacionales. Debe entender que una correlación entre dos variables no necesariamente implica que también existe una relación de causa y efecto. Debe conocer la importancia de un buen muestreo, además, reconocer que muchas encuestas y sondeos de opinión obtienen muy buenos resultados, aunque los tamaños de la muestra puedan parecer relativamente pequeños. Aunque muchas personas rechazan creerlas, una encuesta a nivel nacional de sólo 1 700 votantes puede proporcionar buenos resultados si el muestreo es planificado y realizado con cuidado.

Hubo una época en que una persona se consideraba educada si sabía leer, pero estamos en un nuevo milenio que es mucho más demandante. En la actualidad, una persona educada debe ser capaz de leer, escribir y entender el significado del Renacimiento, operar una computadora y aplicar razonamiento estadístico. El estudio de estadística nos ayuda a ver verdades que en ocasiones están distorsionadas por una falla al enfocar un problema cuidadosamente o están ocultas por datos desorganizados. La comprensión de estadística ahora es esencial tanto para empleados como para empleadores, para todos los ciudadanos. H. G. Wells una vez dijo: “el pensamiento estadístico un día será tan necesario para la eficiencia ciudadana como la capacidad para leer y escribir”. Ese día es ahora.

Apéndice A: tablas de puntuación z



Esta tabla es una versión más detallada de la tabla 5.1; observe que las áreas debajo de la curva mostrada aquí corresponden a *percentiles*. Para leer esta tabla, encuentre los primeros dos dígitos de la puntuación *z* en la columna izquierda. Las puntuaciones *z* negativas están en la página izquierda y las puntuaciones *z* positivas en la página derecha.

Tabla A-1 Distribución normal estándar (z): área acumulada del lado IZQUIERDO										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–3.50 y menores	.0001									
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Observación: para valores de *z* menores a –3.49, utilice 0.0001 para el área.

*Utilice estos valores comunes que resultan de interpolación:

Puntuación <i>z</i>	Área
–1.645	0.0500
–2.575	0.0050

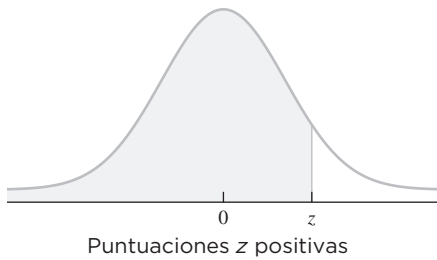


Tabla A-1 (continuación) Área acumulada del lado IZQUIERDO

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50 y mayores	.9999									

Observación: para valores de z mayores a 3.49, utilice 0.9999 para el área.

*Utilice estos valores comunes que resultan de interpolación:

Puntuación z	Área
1.645	0.9500
2.575	0.9950

Tabla A-2 Valores críticos de z			
	Prueba de cola izquierda	Prueba de cola derecha	Prueba de dos colas
Nivel de significancia 0.05	-1.645	1.645	-1.96 y 1.96
Nivel de significancia 0.01	-2.33	2.33	-2.576 y 2.576

Apéndice B: tabla de números aleatorios

Para muchas aplicaciones estadísticas es útil generar un conjunto de números elegidos de manera aleatoria. Puede generar tales números en la mayoría de las calculadoras y computadoras, pero en ocasiones es más sencillo utilizar una tabla como la que se muestra a continuación. Esta tabla fue generada mediante la selección aleatoria de uno de los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8 o 9 para cada posición en la tabla; esto es, cada uno de estos dígitos es igualmente probable que aparezca en cualquier posición. Así, usted puede generar una sucesión de dígitos aleatorios con sólo iniciar en cualquier punto en la tabla y tomar los dígitos en el orden en el que aparezcan. Un conjunto más grande de números aleatorios está disponible en el sitio web (www.aw.com/bbt).

Ejemplo 1: genere una lista aleatoria de respuestas sí/no.

Solución: inicie en un punto arbitrario de la tabla. Si el dígito es 0, 1, 2, 3 o 4 lo consideramos una respuesta *sí*. Si el dígito es 5, 6, 7, 8 o 9 lo consideramos una respuesta *no*. Continúe a lo largo de la tabla desde su punto inicial, usando cada dígito mostrado para determinar una respuesta *sí* o *no* para su lista.

Ejemplo 2: genere una lista aleatoria de calificación con letra A, B, C, D o F.

Solución: Sea 0 o 1 una A, 2 o 3 una B, 4 o 5 una C, 6 o 7 una D y 8 o 9 una F. Inicie en un punto arbitrario de la tabla y use cada dígito mostrado como una calificación para su lista.

9	9	3	2	7	5	6	0	8	1	6	0	2	3	2	8	3	3	1	2	4	7	6	3	4
9	7	1	8	1	6	6	7	6	6	5	4	4	7	7	6	8	1	7	1	0	8	4	9	9
8	1	7	5	0	7	8	5	2	0	9	4	3	9	0	7	6	1	9	1	0	8	7	3	4
0	5	1	0	2	8	7	4	0	3	9	2	6	2	5	8	4	2	2	5	1	9	8	3	4
2	7	8	0	8	1	8	5	6	6	4	4	5	5	4	9	3	5	2	8	6	5	5	4	3
4	8	8	3	3	8	4	6	9	1	8	2	5	7	6	9	7	1	2	3	6	5	1	8	2
5	4	5	8	7	3	8	4	5	7	4	5	2	2	2	7	7	0	2	3	0	2	4	8	6
4	1	9	4	3	0	7	1	9	0	7	3	1	4	0	8	3	2	8	0	5	0	1	0	1
2	8	7	4	6	5	7	7	6	0	0	8	9	5	5	4	0	7	3	9	1	6	3	3	2
5	8	6	8	6	9	6	7	7	5	1	3	5	2	9	7	6	6	3	5	9	4	6	0	5
8	0	9	4	8	5	0	5	6	9	1	0	6	9	5	9	0	7	8	9	9	4	8	3	7
6	1	0	4	1	7	4	0	0	3	5	6	4	2	1	5	1	1	9	0	0	2	5	0	7
3	4	0	9	1	9	3	9	5	1	0	7	4	8	1	1	9	7	0	7	1	4	5	2	6
9	9	6	5	0	7	8	6	1	0	4	9	8	7	7	4	4	7	4	0	7	8	6	4	9
1	5	0	7	8	1	6	3	6	9	5	4	9	5	4	2	4	6	0	4	3	2	6	8	4
5	7	2	8	5	8	3	1	6	4	4	2	2	3	7	0	6	6	3	2	3	5	0	4	6
1	2	3	5	7	1	9	7	7	6	5	5	3	1	9	6	0	5	1	6	3	8	0	3	4
3	4	9	3	5	0	3	7	4	6	2	1	8	5	1	4	1	7	0	2	1	4	9	4	9
2	4	2	6	6	9	9	7	2	1	7	7	3	2	0	6	7	0	7	3	9	3	3	7	5
9	9	1	5	2	8	3	9	9	4	8	9	6	0	6	7	4	6	3	1	3	6	9	8	3
2	6	3	3	9	3	4	1	9	3	0	6	9	1	2	4	1	9	9	8	7	3	2	6	9
1	3	7	8	0	1	0	6	1	6	5	4	9	3	0	7	0	7	2	3	1	7	8	9	2
8	8	4	8	4	0	2	8	5	5	7	3	7	1	2	1	8	3	5	2	0	1	5	3	2
4	2	9	2	8	8	8	9	1	2	4	6	7	1	7	5	1	5	1	9	3	2	2	8	0
1	3	7	2	1	0	6	4	7	6	1	0	8	4	8	9	7	6	3	5	5	1	2	2	9

Lecturas sugeridas

Against the Gods: The Remarkable Story of Risk, P. Bernstein, Wiley, 1996.

All the Math That's Fit to Print: Articles from the Manchester Guardian, K. Devlin, Mathematical Association of America, 1994.

The Arithmetic of Life, G. Shaffner, Ballantine Books, 1999.

The Arithmetic of Life and Death, G. Shaffner, Ballantine Books, 2001.

The Bell Curve Debate, R. Jacoby y N. Glauber (eds.), Times Books, 1995.

Beyond Numeracy: Ruminations of a Number Man, J. A. Paulos, Vintage, 1992.

Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future, D. H. Meadows, D. L. Meadows, y J. Randers, Chelsea Green Publishing Company, 1992.

Billions and Billions, C. Sagan, Random House, 1997.

The Broken Dice and Other Mathematical Tales of Chance, I. Ekeland, University of Chicago Press, 1993.

Can You Win?, M. Orkin, Freeman, 1991.

The Cartoon Guide to Statistics, L. Gonick y W. Smith, HarperCollins, 1993.

The Complete How to Figure It, D. Huff, Norton, 1996.

Damned Lies and Statistics, J. Best, University of California Press, 2001.

Ecological Numeracy: Quantitative Analysis of Environmental Issues, R. Herendeen, Wiley, 1998.

Elementary Statistics, décima edición, M. Triola, Addison-Wesley, 2006.

Emblems of Mind: The Inner Life of Music and Mathematics, E. Rothstein, Random House, 1995.

Envisioning Information, E. Tufte, Connecticut Graphics Press, Cheshire, 1983.

Flaws and Fallacies in Statistical Thinking, S. Campbell, Prentice-Hall, 1974.

Games, Gods, and Gambling, F. N. David, Dover, 1998.

Go Figure! The Numbers You Need for Everyday Life, N. Hopkins y J. Mayne, Visible Ink Press, 1992.

The Golden Mean, C. F. Linn, Doubleday, 1974.

The Honest Truth About Lying with Statistics, C. Holmes, Charles C. Thomas, 1990.

How Many People Can the Earth Support, J. E. Cohen, Norton, 1995.

How to Lie with Statistics, D. Huff, Norton, 1993.

How to Tell the Liars from the Statisticians, R. Hooke, Dekker, 1983.

How to Use (and Misuse) Statistics, G. Kimble, Prentice-Hall, 1978.

Innumeracy, J. A. Paulos, Hill y Wang, 1988.

The Jungles of Randomness, I. Peterson, Wiley, 1998.

Lady Luck: The Theory of Probability, W. Weaver, Anchor Books, 1963.

The Lady Tasting Tea: How Statistics Revolutionized Science in the 20th Century, D. Salsburg, Freeman, 2001.

Life by the Numbers, K. Devlin, Wiley, 1998.

The Mathematical Tourist, I. Peterson, Freeman, 1988.

A Mathematician Reads the Newspaper, J. A. Paulos, Basic Books, 1995.

Mathematics, the Science of Patterns: The Search for Order in Life, Mind, and the Universe, K. Devlin, Scientific American Library, 1994.

The Mismeasure of Man, S. J. Gould, Norton, 1981.

The Mismeasure of Woman, C. Tavis, Touchstone Books, 1993.

Misused Statistics: Straight Talk for Twisted Numbers, A. Jaffe y H. Spier, Dekker, 1987.

Nature's Numbers, I. Stewart, Basic Books, 1995.

Number: The Language of Science, T. Dantzig, Macmillan, 1930.

Once Upon a Number, J. A. Paulos, Basic Books, 1998.

Overcoming Math Anxiety, S. Tobias, Houghton Mifflin, 1978. Edición revisada, Norton, 1993.

Pi in the Sky: Counting, Thinking, and Being, J. D. Barrow, Little, Brown, 1992.

The Population Explosion, P. R. Ehrlich y A. H. Ehrlich, Simon y Schuster, 1990.

Probabilities in Everyday Life, J. McGervey, Ivy Books, 1989.

The Psychology of Judgment and Decision Making, S. Plous, McGraw-Hill, 1993.

Randomness, D. Bennett, Harvard University Press, 1998.

The Statistical Exorcist: Dispelling Statistics Anxiety, M. Hollander y F. Proschan, Dekker, 1984.

Statistics, tercera edición, D. Freedman, R. Pisani, R. Purves, y A. Adhikari, Norton, 1997.

Statistics: Concepts and Controversies, quinta edición, D. Moore, Freeman, 2001.

Statistics with a Sense of Humor, F. Pyrczak, Fred Pyrczak Publisher, 1989.

Tainted Truth: The Manipulation of Fact in America, C. Crossen, Simon y Schuster, 1994.

The Tipping Point: How Little Things Can Make a Big Difference, M. Gladwell, Little, Brown, 2000.

200% of Nothing, A. K. Dewdney, Wiley, 1993.

The Universe and the Teacup: The Mathematics of Truth and Beauty, K. C. Cole, Harcourt Brace, 1998.

The Visual Display of Quantitative Information, E. Tufte, Connecticut Graphics Press, Cheshire, 1983.

Vital Signs, compilada por the Worldwatch Institute, Norton, 1998 (actualizada cada año).

What Are the Odds? Chance in Everyday Life, L. Krantz, Harper Perennial, 1992.

Créditos

CRÉDITOS DE FIGURAS Y TEXTO

Capítulo 1 Figura 1.1: *New York Times*, 1 de noviembre de 1999. Copyright © 1999. The New York Times Co. Reimpreso con permiso.

Capítulo 2 Figura 2.4: *New York Times*, 3 de enero de 2007. Copyright © 2007 The New York Times Co. Reimpreso con permiso.

Capítulo 3 Figura 3.15: partes de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 540. © 1999 Dow Jones & Co. Figura 3.16: de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 694. © 1999 Dow Jones & Co. Reimpreso con permiso de Ballantine Books, una marca de Random House, Inc. Figura 3.18: de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 238. © 1999 Dow Jones & Co. Reimpreso con permiso de Ballantine Books, una marca de Random House, Inc. Figura 3.19: *New York Times*, 30 de septiembre de 1995. Copyright © 1995 The New York Times Co. Reimpreso con permiso. Figura 3.22: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 365. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.24: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 367. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.25: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 377. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.27: *New York Times*, 3 de octubre de 1995. Copyright © 1995 The New York Times Co. Reimpreso con permiso. Figura 3.28: partes de *New York Times*, 20 de agosto de 2000. Copyright © 2000 The New York Times Co. Reimpreso con permiso. Figura 3.33: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 376. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.34: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 375. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.35: de Bennett/Briggs, *Using and Understanding Mathematics*, segunda edición, página 330. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.36: *New York Times*, 2 de abril de 2000. Copyright © 2000 The New York Times Co. Reimpreso con permiso. Figura 3.40: partes de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 323. © 1999 Dow Jones & Co. Figura 3.41: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 371. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.42: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 371. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.43: de Bennett et al., *The Cosmic Perspective*, cuarta edición, página 378. Copyright © 2007 Pearson Education, Inc. Reimpreso con permiso. Figura 3.34: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 372. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 3.50: de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 577. © 1999 Dow Jones & Co. Figura 3.52: de Mortensen, Pedersen, Westergaard, Wohlfahrt, Ewald, Mors, Andersen, Melbye. "Effects of Family History and Place and Season of Birth on the Risk of Schizophrenia", *The New England Journal of Medicine*, 25/feb/99, página 606. Copyright © 1999 Sociedad Medica de Massachusetts. Todos los derechos reservados.

Figura 3.54: de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 545. © 1999 Dow Jones & Co. Figura 3.59: adaptado de Edward R. Tufte, *The Visual Display of Quantitative Information*, Cheshire, CT: Graphics Press, 1983. Reimpreso con permiso. Figura 3.58: partes de *The Wall Street Journal Almanac* de los editores de *The Wall Street Journal*, 1999, página 662. © 1999 Dow Jones & Co. Figura 3.60: de Bennett et al., *Essential Cosmic Perspective*, cuarta edición, página 216. Copyright © 2009 Pearson Education, Inc. Reimpreso con permiso. Figura 3.61: de Bennett et al., *The Essential Cosmic Perspective*, cuarta edición, página 217. Copyright © 2009 Pearson Education, Inc. Reimpreso con permiso.

Capítulo 4 Figura 4.1: de Mario F. Triola, *Elementary Statistics*, séptima edición, página 61. © 1998 Addison Wesley Longman Inc. Reimpreso con permiso de Pearson Education. Figura 4.9: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 411. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 4.10: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 411. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 4.11: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 411. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 4.17: de Bennett/Briggs, *Using and Understanding Mathematics*, cuarta edición, página 258. Copyright © 2008 Pearson Education, Inc. Reimpreso con permiso de Pearson Education, Inc.

Capítulo 5 Figura 5.26: de Mario F. Triola, *Elementary Statistics*, séptima edición, página 256. © 1998 Addison Wesley Longman Inc. Reimpreso con permiso de Pearson Education. Figura 5.28: *New York Times*, 24 de febrero de 1998. Copyright © 1998 The New York Times Co. Reimpreso con permiso.

Capítulo 6 Epígrafe: Ogden Nash Poems & Stories de Linell Nash Smith e Isabel Nash Eberstadt. Figura 6.5: de *The Virginia Pilot*, 14/sept/99. Figura 6.13: partes de *New York Times*, 11 de diciembre de 1999. Copyright © 1999 The New York Times Co. Reimpreso con permiso. Figura 6.19: de *Statistical Science* de Kathryn Roeder, 15 de mayo de 1994. Copyright 1994 del Instituto de Estadística Médica. Todos los derechos reservados. Reimpreso con permiso por medio de los derechos reservados de Centro Clearance.

Capítulo 7 Figura 7.21: de Bennett/Briggs, *Using and Understanding Mathematics*, página 501. Copyright © 1999 Addison Wesley Longman, Inc. Reimpreso con permiso de Pearson Education, Inc. Figura 7.25: de Bennett et al., *The Cosmic Perspective*, cuarta edición, página 293. Copyright © 2007 Pearson Education, Inc. Reimpreso con permiso. Figura 7.26: de Bennett et al., *The Cosmic Perspective*, cuarta edición, página 321. Copyright © 2007 Pearson Education, Inc. Reimpreso con permiso.

Capítulo 9 Figura 9.10: *New York Times*, 3 de enero de 2007. Copyright © 2007 The New York Times Co. Reimpreso con permiso. Figura 9.11: *New York Times*, 3 de enero de 2007. Copyright © 2007 The New York Times Co. Reimpreso con permiso.

Capítulo 10 Figura 10.4: *New York Times*, 4 de agosto de 1998. Copyright © 1998 The New York Times Co. Reimpreso con permiso. Figura 10.5: *New York Times*, 4 de agosto de 1998, página B10. Copyright © 1998 The New York Times Co. Reimpreso con permiso.

CRÉDITOS DE FOTOGRAFÍAS

Capítulo 1 Página 1, © Time & Life Pictures/Getty/Images; página 3, © Getty Images Sport; página 6, PhotoDisc; página 8, © AFP/Corbis; página 12, © Taxi/Getty Images; página 14, cortesía de Jeff Bennett; página 16, PhotoDisc; página 25, Corbis Royalty Free; página 28, © Digital Vision; página 30, PhotoDisc; página 35, PhotoDisc; página 38, NASA; página 41, © Michael J. Okoniewski, 143 Avenida Peck, Syracuse NY, 13206; página 48, PhotoDisc; página 50, © Digital Vision

Capítulo 2 Página 53, PhotoDisc; página 60, Corbis Royalty Free; página 63, NASA; página 64, PhotoDisc; página 70, © Peter Turnley/Corbis; página 71, © Digital Vision; página 75, © Gene Blevins/Corbis; página 78, © Bettmann/Corbis; página 84, Corbis Royalty Free; página 86, Image Source/Getty Royalty Free

Capítulo 3 Página 89, © AFP/Getty Images; página 92, Corbis Royalty Free; página 95, © Stone/Getty Images; página 103, © Ralf-Finn Hestoft/Corbis; página 106, © Everett Collection; página 125, © AP Wideworld Photo; página 126, © PhotoDisc Red; página 128, © Charles & Josette Lenars/Corbis; página 137, © Archivo Iconográfico, S. A./Corbis; página 140, PhotoDisc

Capítulo 4 Página 145, PhotoDisc; página 149, © PCN Photography; página 161, © Laurent Rebours/AP Wideworld Photos; página 168, fotografía de Beth Anderson; página 173, © Toru Hanai/Reuters/Corbis; página 181, © Richard T. Nowitz/Corbis; página 188, © John-Marshall Mantel/Corbis; página 191, © James L. Amos/Corbis

Capítulo 5 Página 195, (izquierda) © Dimitri Lundt, TempSport/Corbis; (derecha) © Getty Images Sport; página 198, fotografía de Beth Anderson; página 206, PhotoDisc; página 210, Armada de Estados Unidos; página 226, fotografía de Beth Anderson; página 229, PhotoDisc

Capítulo 6 Página 233, PhotoDisc; página 240, (superior) PhotoDisc; (inferior) PhotoDisc; página 241, PhotoDisc; página 245, Gary Holscher/Stone/Getty Images; página 252, © Getty Images News; página 260, Stockbyte/Getty Royalty Free; página 268, © Greg Fiume/NewSport/Corbis; página 270, © Bettmann/Corbis; página 279, Photographer's Choice Getty Royalty Free; página 282, Corbis Royalty Free

Capítulo 7 Página 285, PhotoDisc; página 287, PhotoDisc; página 300, PhotoDisc; página 303, © Tomasso DeRosa/Corbis; página 307, PhotoDisc Red; página 313, StockDisc CD SD174; página 317, PhotoDisc; página 318, © AP Wideworld Photos; página 325, Blend Images/Getty Royalty Free; página 328, © Digital Vision

Capítulo 8 Página 333, © Digital Vision; página 340, © James A Sugar/Corbis; página 346, © Digital Vision; página 363, PhotoDisc

Capítulo 9 Página 369, PhotoDisc; página 370, PhotoDisc; página 377, © Digital Vision; página 390, © Layne Kennedy/Corbis; página 403, © Digital Vision; página 406, PhotoDisc CD Vol 12

Capítulo 10 Página 409, © Photonica/Getty Images; página 439, fotografía de Beth Anderson; página 443, fotografía de Beth Anderson

Glosario

aleatorización El proceso de asegurar que los sujetos de un experimento son asignados al grupo de tratamiento o al grupo de control al azar y de tal manera que cada sujeto tenga igual oportunidad de ser asignado a cualquier grupo.

análisis de varianza (ANOVA) Un método para probar la igualdad de tres o más medias poblacionales analizando sus varianzas muestrales.

ANOVA Vea *análisis de varianza*.

barras Un diagrama que consiste en barras que representan las frecuencias (o frecuencias relativas) para categorías particulares. Las longitudes de las barras son proporcionales a las frecuencias.

cambio absoluto El aumento o disminución real de una valor de referencia a un valor nuevo:

$$\text{cambio absoluto} = \text{valor nuevo} - \text{valor de referencia}$$

cambio relativo El tamaño de un cambio absoluto en comparación al valor de referencia, expresado como un porcentaje:

$$\text{cambio relativo} = \frac{\text{valor nuevo} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\%$$

causalidad Una relación presente cuando una variable es una causa de otra.

censo La colección de datos de cada miembro de una población.

ciego La práctica de mantener a los sujetos experimentales y/o a los experimentadores en la ignorancia acerca de quién está en el grupo de tratamiento y quién está en el grupo de control. También vea *experimento ciego sencillo*; *experimento doble ciego*.

clase mediana Para datos agrupados, la clase en la que se encuentra el valor mediano.

coeficiente de correlación (r) Una medida de la fuerza de la relación entre dos variables. Su valor siempre está entre -1 y 1 (es decir, $-1 \leq r \leq 1$).

coeficiente de determinación (R^2) Un número que describe cuán bien se ajustan los datos a una ecuación de mejor ajuste mediante regresión múltiple.

complemento Para un evento A , todos los resultados en los que A no ocurre, se expresa como $\bar{A}$. Su probabilidad es $P(\bar{A}) = 1 - P(A)$.

confusión Error en la interpretación de resultados estadísticos, que ocurre cuando los efectos de diferentes factores se mezclan tal que los efectos de los factores individuales que son estudiados no pueden determinarse.

correlación Una relación estadística entre dos variables. También vea *correlación negativa*; *correlación positiva*; *sin correlación*.

correlación negativa Una correlación en la cual las dos variables tienden a cambiar en direcciones opuestas, una aumenta mientras la otra disminuye.

correlación positiva Un tipo de correlación en el que dos variables tienden a aumentar (o disminuir) juntas.

cuartil Valor que divide a una distribución de datos en cuatro partes iguales.

cuartil central La mediana global de un conjunto de datos. También llamado *segundo cuartil*.

cuartil inferior La mediana de los datos en la mitad inferior de un conjunto de datos. También conocido como *primer cuartil*.

cuartil superior La mediana de los datos en la mitad superior de un conjunto de datos. También conocido como *tercer cuartil*.

dato extremo Un valor en un conjunto de datos que es mucho mayor o mucho menor que casi todos los demás valores.

datos continuos Datos cuantitativos que pueden tomar cualquier valor en un intervalo dado.

datos cualitativos Datos que consisten en valores que describen cualidades o categorías no numéricas.

datos cuantitativos Datos que consisten en valores que representan conteos o medidas. Los datos cuantitativos pueden ser discretos o continuos.

datos discretos Datos cuantitativos que pueden tomar sólo valores particulares y ningún otro en medio (por ejemplo, los números enteros 0, 1, 2, 3, 4, 5).

datos geográficos Datos que pueden asignarse a diferentes ubicaciones geográficas.

datos sin procesar Las mediciones u observaciones reales recolectadas de una muestra.

desviación Cuán lejos el valor de un dato particular está de la media de un conjunto de datos, usada para calcular la desviación estándar.

desviación estándar Un solo número que comúnmente es utilizado para describir la variación en una distribución de datos, calculada como

$$\text{desviación estándar} = \sqrt{\frac{\text{suma de todas (desviaciones de la media)}^2}{\text{número total de valores de datos} - 1}}$$

diagrama apilado Un tipo de gráfica de barras o diagrama de líneas en el que dos o más conjuntos diferentes de datos son apilados verticalmente.

diagrama circular (de pastel) Un círculo dividido de modo que cada sector representa la frecuencia relativa de una categoría particular. El tamaño del sector es proporcional a la frecuencia relativa y el total del círculo representa la frecuencia relativa total de 100%.

diagrama de caja Una exhibición gráfica de un resumen de cinco números. Una línea numérica se utiliza como referencia, los valores de los cuartiles inferior y superior son encerrados en una caja, una recta se dibuja a través de la caja para la mediana, y dos “bigotes” se extienden para los valores inferiores y superiores. También llamado *diagrama de caja y bigotes*.

diagrama de dispersión Una gráfica, con frecuencia usada para investigar correlaciones, en la que cada punto corresponde a los valores de dos variables. También llamado *gráfica de dispersión*.

diagrama de múltiples líneas Una extensión sencilla de una gráfica de líneas común, en la que dos o más líneas permiten la comparación de dos o más conjuntos de datos.

diagrama de Pareto Una gráfica de barras en el nivel nominal de medida, con las barras acomodadas en orden de frecuencia.

diagrama de serie de tiempo Un histograma o diagrama de líneas en el que el eje horizontal representa tiempo.

diagrama de tallo y hojas Una gráfica que se parece mucho a un histograma girado, con listas de datos individuales en lugar de barras. También llamado *diagrama de tallos*.

diferencia absoluta La diferencia real entre el valor comparado y el valor de referencia:

$$\text{diferencia absoluta} = \text{valor comparado} - \text{valor de referencia}$$

diferencia relativa El tamaño de una diferencia absoluta en comparación al valor de referencia, expresado como un porcentaje:

$$\text{diferencia relativa} = \frac{\text{valor comparado} - \text{valor de referencia}}{\text{valor de referencia}} \times 100\%$$

distribución La forma en que los valores de una variable están dispersos sobre todos los valores posibles. Puede mostrarse mediante una tabla o con una gráfica.

distribución binomial Una distribución con dos picos o modas.

distribución con un solo pico Una distribución con una sola moda. También vea *distribución unimodal*.

distribución de medias muestrales La distribución que resulta cuando se encuentran las medias ($\bar{x}$) de todas las muestras posibles de un tamaño dado.

distribución de probabilidad La distribución completa de las probabilidades de todos los eventos posibles asociados con una variable particular. Puede mostrarse como una tabla o como una gráfica.

distribución de proporciones muestrales La distribución que resulta cuando se encuentran las proporciones ($\hat{p}$) de todas las muestras posibles de un tamaño dado.

distribución muestral La distribución de un estadístico muestral, tal como una media o proporción, tomada de todas las posibles muestras de un tamaño particular.

distribución normal Un tipo especial de distribución simétrica, en forma de campana con un solo pico que corresponde a la media, mediana y moda de la distribución. Su variación puede caracterizarse mediante la desviación estándar. También vea *regla 68-95-99.7*.

distribución sesgada a la derecha Una distribución en la cual los valores están más esparcidos en el lado derecho.

distribución sesgada a la izquierda Una distribución en la cual los valores están más dispersos en el lado izquierdo.

distribución simétrica Una distribución en la cual la mitad izquierda es un espejo de la mitad derecha.

distribución t Una distribución que es muy semejante en forma y simetría a la distribución normal pero que toma en cuenta la gran variabilidad esperada con muestras pequeñas. Tiende a la distribución normal para tamaños grandes de muestras.

distribución uniforme Una distribución en la que todos los valores tienen la misma frecuencia.

distribución unimodal Vea *distribución con un solo pico*.

efecto del experimentador Un efecto que ocurre cuando un investigador o experimentador influye de alguna manera en los sujetos a través de factores como expresiones faciales, tono de voz o actitud.

efecto placebo Un efecto en el que los pacientes mejoran sólo porque ellos creen que están recibiendo un tratamiento útil, cuando en realidad ellos podrían estar recibiendo sólo un placebo.

encuesta de autoselección Una encuesta en la cual la persona decide por ella misma si será incluida. También denominada *encuesta de respuesta voluntaria*.

encuesta de respuesta voluntaria Vea *encuesta de autoselección*.

error absoluto La cantidad real en que difiere el valor que se mide y el valor verdadero:

$$\text{error absoluto} = \text{valor medido} - \text{valor verdadero}$$

error de tipo I En una prueba de hipótesis, el error de rechazar la hipótesis nula, H_0 , cuando es verdadera.

error de tipo II En una prueba de hipótesis, el error de no rechazar la hipótesis nula, H_0 , cuando es falsa.

error muestral Error introducido cuando una muestra aleatoria se utiliza para estimar un parámetro poblacional; la diferencia entre un resultado muestral y un parámetro poblacional.

error relativo La cantidad relativa por la cual un valor medido difiere del valor verdadero, expresado como un porcentaje:

$$\text{error relativo} = \frac{\text{valor medido} - \text{valor verdadero}}{\text{valor verdadero}} \times 100\%$$

errores aleatorios Errores que ocurren a consecuencia del azar y de eventos inherentemente impredecibles en el proceso de medición.

errores sistemáticos Errores que ocurren cuando existe un problema en el sistema de medición que afecta todas las mediciones de la misma manera.

esperanza de vida El número de años que una persona de una edad dada hoy puede esperarse que viva, en promedio. Tiene como base estadísticas médicas y de salud actuales y no toma en cuenta cambios futuros en la ciencia médica o salud pública.

estadística La ciencia que recolecta, organiza e interpreta datos.

estadísticas Los datos que describen o resumen algo.

estadísticas vitales Datos concernientes a nacimientos y muertes de personas.

estadístico ji cuadrada (χ^2) Un número usado para determinar la significancia estadística de una prueba de hipótesis en una tabla de contingencia (o tabla de dos entradas). Si es menor que un valor crítico (que depende del tamaño de la tabla y el nivel de significancia), las diferencias entre las frecuencias observadas y las frecuencias esperadas no son significativas.

estadísticos muestrales Características de la muestra que se encuentran mediante la consolidación o resumen de datos sin procesar.

estudio de control de caso Un estudio de observación que parece un experimento, ya que la muestra se divide de manera natural en dos (o más) grupos. Los participantes con el comportamiento bajo estudio forman los casos, como el grupo de tratamiento en un experimento. Los participantes que sin el comportamiento son los *controles*, como el grupo de control en un experimento.

estudio observacional Un estudio en el que los investigadores observan o miden características de los elementos de la muestra, pero no intentan influir o modificar estas características.

evento En probabilidad, una colección de uno o más resultados que comparten una propiedad de interés. También vea *resultado*.

eventos dependientes Dos eventos para los que el resultado de uno afecta la probabilidad del otro.

eventos independientes Dos eventos para los cuales un resultado de uno no afecta la probabilidad del otro.

eventos que no se traslapan Dos eventos para los cuales la ocurrencia de uno excluye la ocurrencia del otro.

eventos que se traslapan Dos eventos que posiblemente puedan ocurrir ambos.

experimento Un estudio en el que los investigadores aplican un tratamiento y luego observan su efecto en los sujetos.

experimento ciego sencillo Un experimento en el que los participantes no saben si son miembros del grupo de tratamiento o del grupo de control, pero los experimentadores sí saben o los participantes saben pero los experimentadores no.

experimento doble ciego Un experimento en el que ni los participantes ni los experimentadores saben quién pertenece al grupo de tratamiento y quién pertenece al grupo de control.

factores de confusión Cualesquiera factores o variables en un estudio estadístico que puede tender a confundir. También conocido como *variables de confusión*.

falacia del jugador La creencia errónea de que una racha de mala suerte provoca que una persona “deba” tener una racha de buena suerte.

frecuencia Para una categoría de datos, el número de veces que valores de datos caen en esa categoría.

frecuencia acumulada Para cualquier categoría de datos, el número de valores de datos en esa categoría y de todas las categorías precedentes.

frecuencia esperada En una tabla de dos entradas, la frecuencia que uno espera en una celda dada de la tabla, si las variables de renglón y de columna fuesen independientes entre sí.

frecuencia relativa Para cualquier categoría de datos, la fracción o porcentaje de la frecuencia total que cae en esa categoría:

$$\text{frecuencia relativa} = \frac{\text{frecuencia en la categoría}}{\text{frecuencia total}}$$

grados de libertad (para una distribución t) El tamaño de la muestra menos uno ($n - 1$).

gráfica de clasificación Agrupación de datos en categorías (clases), cada una de las cuales cubre un rango de posibles valores de datos.

gráfica de líneas Una gráfica que muestra la distribución de datos cuantitativos como una serie de puntos conectados mediante líneas. La posición horizontal de cada punto corresponde al centro de la clase que representa y la posición vertical corresponde al valor de la frecuencia para la clase.

gráfica de múltiples barras Una extensión sencilla de una gráfica de barras común, en la que dos o más conjuntos de barras permiten la comparación de dos o más conjuntos de datos.

gráfica de puntos Un diagrama semejante a una gráfica de barras excepto que cada valor individual está representado por un punto.

grupo de control El grupo de sujetos en un experimento que no reciben el tratamiento que está siendo probado.

grupo de tratamiento El grupo de sujetos en un experimento que reciben el tratamiento que será probado.

hipótesis En estadística, una afirmación acerca de un parámetro poblacional, tal como una proporción poblacional, p , o una media poblacional, μ . También vea *hipótesis alternativa*; *hipótesis nula*.

hipótesis alternativa (H_a) Un enunciado que puede aceptarse sólo si la hipótesis nula es rechazada.

hipótesis nula (H_0) Una afirmación específica (tal como un valor específico para un parámetro poblacional) contra la cual se prueba una hipótesis alternativa.

histograma Una gráfica de barras que muestra una distribución para datos cuantitativos (en el nivel de medida de intervalo o nivel de razón). Las barras tienen un orden natural y los anchos de las barras tienen significado específico.

Índice de precios al consumidor (IPC) Un número índice diseñado para medir la tasa de inflación. Es calculado y reportado mensualmente, con base en una muestra de más de 60 000 costos de bienes, servicios y vivienda.

inflación El aumento, a lo largo del tiempo, de precios y salarios. Su tasa global es medida mediante el IPC.

intervalo de confianza Un rango de valores asociados con un nivel de confianza, tal como 95%, que es probable que contenga el valor verdadero de un parámetro poblacional.

ley de los grandes números Un resultado importante en probabilidad que se aplica a un proceso para el cual la probabilidad de un evento A es $P(A)$ y los resultados de ensayos repetidos son independientes. Establece: si el proceso es repetido muchas veces, entre mayor sea el número de ensayos, la proporción debe ser más cercana a $P(A)$. También conocida como *ley de promedios*.

mapa de contornos Un mapa que utiliza curvas (contornos) para conectar regiones geográficas con los mismos valores de datos.

margen de error La máxima diferencia probable entre un estadístico muestral observado y el valor verdadero de un parámetro poblacional. Su tamaño depende del nivel de confianza deseado.

media La suma de todos los valores divididos entre el número total de valores. Es lo que se conoce como el valor promedio.

media muestral La media de una muestra, denotada $\bar{x}$ (“x barra”).

media poblacional La media verdadera de una población, denotada por la letra griega μ (se pronuncia “mu”).

media ponderada Una media que toma en cuenta diferencias en la importancia relativa de los datos. A cada dato se le asigna un peso y entonces

$$\text{media ponderada} = \frac{\text{suma de (cada dato} \times \text{su peso)}}{\text{suma de todos los pesos}}$$

mediana El valor de la mitad en un conjunto ordenado de datos (o la mitad entre los dos valores de en medio, si el número de valores es par).

metaanálisis Un estudio en el que los investigadores analizan muchos estudios individuales (sobre un tema particular) como un grupo combinado, con el objetivo de encontrar tendencias que no fueron evidentes en los estudios individuales.

método a priori Vea *método teórico*.

método de la frecuencia relativa Un método de estimación de una probabilidad con base en las observaciones o experimentos usando la frecuencia observada o frecuencia relativa medida del evento de interés. También conocido como *método empírico*.

método empírico Vea *método de la frecuencia relativa*.

método subjetivo Un método de estimación de una probabilidad con base en la experiencia o la intuición.

método teórico Un método para estimar una probabilidad con base en una teoría, o conjunto de supuestos, acerca del proceso en cuestión. Suponiendo que todos los resultados son igualmente probables, la probabilidad teórica de un evento particular se encuentra dividiendo el número de formas en que

el evento ocurre entre el número total de posibles resultados. También denominado *método a priori*.

métodos de muestreo Vea *muestreo aleatorio simple*; *muestreo de conveniencia*; *muestreo estratificado*; *muestreo por conglomerados*; *muestreo sistemático*.

moda El valor más común (o grupo de valores) en una distribución.

muestra Un subconjunto de la población de la cual los datos son obtenidos.

muestra representativa Una muestra en la cual las características relevantes de los elementos son generalmente las mismas que las características de la población.

muestreo El proceso de elección de una muestra de una población.

muestreo aleatorio simple Una muestra de elementos elegidos de tal manera que cada posible muestra del mismo tamaño tenga una igual oportunidad de ser elegida.

muestreo de conveniencia Selección de una muestra que está disponible fácilmente.

muestreo estratificado Un método de muestreo que toma en cuenta las diferencias entre subgrupos, o estratos, dentro de una población. Primero los estratos son identificados y luego se saca una muestra aleatoria dentro de cada estrato. La muestra total consiste en todas las muestras de los estratos individuales.

muestreo por conglomerados División de la población en grupos, o conglomerados, seleccionando alguno de estos conglomerados al azar, y luego obteniendo la muestra al seleccionar todos los miembros dentro de cada conglomerado.

muestreo sistemático Uso de un sistema sencillo para elegir la muestra, tal como seleccionar cada 10 o cada 50 elementos de la población.

nivel de medida Vea *nivel de medida de intervalo*; *nivel de medida de razón*; *nivel de medida nominal*; *nivel de medida ordinal*.

nivel de medida de intervalo Un nivel de medida para datos cuantitativos en el que las diferencias, o intervalos, son significativos, pero las razones o cocientes no. Los datos en este nivel tienen un punto inicial arbitrario.

nivel de medida de razón Un nivel de medida para datos cuantitativos en los que tanto intervalos como razones tienen significado. Los datos en este nivel tienen un punto cero verdadero.

nivel de medida nominal Un nivel de medida para datos cualitativos que consiste sólo en nombres, rótulos o categorías y no pueden clasificarse u ordenarse.

nivel de medida ordinal Un nivel de medida para datos cualitativos que pueden acomodarse en algún orden. Por lo general no tiene sentido hacer comparaciones con los datos.

número índice Un número que proporciona una forma sencilla de calcular medidas en diferentes tiempos o diferentes lugares. El valor en un tiempo particular (o lugar) debe elegirse que sea el valor de referencia (o valor base). El número índice para cualquier otro tiempo (o lugar) es

$$\text{número índice} = \frac{\text{valor}}{\text{valor de referencia}} \times 100$$

paradoja de Simpson Una paradoja estadística que surge cuando los resultados de todo un grupo parecen ser inconsistentes con los de sus subgrupos, puede suceder siempre que los subgrupos no sean de igual tamaño.

parámetros poblacionales Características específicas de la población que un estudio estadístico está diseñado para estimar.

participantes Personas (en lugar de objetos) quienes son los sujetos de un estudio.

percentiles Valores que dividen una distribución de datos en 100 segmentos, cada uno representando 1% de los valores de datos.

pictograma Una gráfica embellecida con ilustraciones.

placebo Algo que carece de los ingredientes activos de un tratamiento pero es idéntico en apariencia a éste. Por tanto, los participantes en un estudio no pueden distinguir el placebo del tratamiento real.

población El conjunto completo de personas o cosas que están siendo estudiadas.

precisión Qué tan cerca una medida aproxima a un valor verdadero. Una medida precisa es muy cercana al valor verdadero. La cantidad de detalle en una medición.

probabilidad Para un evento, la verosimilitud de que éste ocurrirá. La probabilidad de un evento, escrito como $P(\text{evento})$, siempre está entre 0 y 1. Una probabilidad de 0 significa que el evento es imposible, y una probabilidad de 1 significa que el evento es seguro. También vea *método de la frecuencia relativa*; *método subjetivo*; *método teórico*.

probabilidad condicional La probabilidad de un evento dada la ocurrencia de otro evento, escrito $P(B \text{ dado } A)$ o bien $P(B|A)$.

probabilidad conjunta Vea *probabilidad “y”*

probabilidad o cualquiera La probabilidad de que *cualquier* evento A o el evento B ocurran. Cómo se calcula depende si los eventos se traslapan o no se traslapan.

probabilidad “y” La probabilidad de que el evento A y el evento B ocurran. La forma de calcularse depende si los eventos son independientes o dependientes. También llamada *probabilidad conjunta*.

proporción muestral La proporción de alguna característica en una muestra, denotada $\hat{p}$ (“p gorro”).

proporción poblacional La proporción verdadera de alguna característica en una población, denotada por p .

prueba de cola derecha Una prueba de hipótesis que incluye probar si un parámetro poblacional está a la derecha (valores mayores) de un valor afirmado.

prueba de cola izquierda Una prueba de hipótesis que incluye una prueba si un parámetro poblacional está a la izquierda (valores menores) de un valor afirmado.

prueba de dos colas Una prueba de hipótesis que incluye si un parámetro poblacional está a cualquiera de los lados de un valor afirmado.

prueba de hipótesis Un procedimiento estándar para probar una afirmación acerca del valor de un parámetro poblacional.

prueba de una cola Vea *prueba de cola derecha*; *prueba de cola izquierda*.

puntuación estándar Para un valor particular, el número de desviaciones estándar (usualmente denotado por z) entre él y la media de la distribución:

$$z = \text{puntuación estándar} = \frac{\text{valor del dato} - \text{media}}{\text{desviación estándar}}$$

También llamado *puntuación z*.

puntuación z Vea *puntuación estándar*.

rango Para una distribución, la diferencia entre los valores más pequeño y más grande.

recta de mejor ajuste La recta en un diagrama de dispersión que está más cerca de los puntos de datos que otras rectas posibles (de acuerdo con una medida estadística estándar de cercanía). También conocida como *recta de regresión*.

recta de regresión Vea *recta de mejor ajuste*.

regla 68-95-99.7 Directriz que establece que, para una distribución normal, alrededor de 68% (en realidad, 68.3%) de los datos caen dentro de 1 desviación estándar de la media, alrededor de 95% (en realidad, 95.4%) de los datos caen dentro de dos desviaciones estándar de la media y alrededor de 99.7% de los datos caen dentro de tres desviaciones estándar de la media.

regla de en comparación con regla más que (menos que) Una regla para comparaciones. Establece: si el valor comparado es *P%* más que el valor de referencia, entonces es $(100 + P)\%$ del valor de referencia. Si el valor comparado es *P%* menor que el valor de referencia, entonces es $(100 - P)\%$ del valor de referencia.

regla de redondeo Para cálculos estadísticos, la práctica de establecer las respuestas con un lugar decimal más de precisión que el encontrado en los datos sin procesar. Por ejemplo, la media de 2, 3 y 5 es 3.3333..., que sería redondeado a 3.3.

regla del evento raro Regla que establece que es adecuado concluir que una suposición dada (tal como la hipótesis nula) probablemente no es correcta, si la probabilidad de un evento particular al menos tan extremo como el evento observado es muy pequeña.

regla empírica para el rango Una directriz que estipula, para un conjunto de datos sin datos extremos, que la desviación estándar es aproximadamente igual a rango/4.

regresión múltiple Una técnica que permite el cálculo de una ecuación de mejor ajuste que representa el mejor ajuste entre una variable (tal como el precio) y una *combinación* de dos o más variables (tal como peso y color).

relación no lineal Una relación entre dos variables que no puede representarse con una ecuación lineal (línea recta).

resultado En probabilidad, el resultado más básico de una observación o de un experimento. Vea también *evento*.

resultado estadísticamente significativo Un resultado en un estudio estadístico que es poco probable que haya ocurrido por azar. Los niveles de significancia estadística más comúnmente usados son el nivel 0.05 (la probabilidad de que el resultado haya ocurrido por el azar es 5% o menos, o menos de 1 en 20) y el nivel 0.01 (la probabilidad de que el resultado haya ocurrido por azar es 1% o menos, o menos de 1 en 100).

resumen de cinco números Una descripción de la variación de una distribución de datos en términos del valor mínimo, cuartil inferior, mediana, cuartil superior y valor máximo.

revisión de pares Un proceso mediante el cual varios expertos en un campo evalúan un reporte de investigación antes de que éste sea publicado.

sesgada Vea *distribución sesgada a la derecha*; *distribución sesgada a la izquierda*.

sesgo En un estudio estadístico, cualquier problema en el diseño o realización del estudio que tienda a favorecer ciertos resultados. Vea también *sesgo de participación*; *sesgo de selección*.

sesgo de participación Sesgo que ocurre cuando la participación en un estudio es voluntaria.

sesgo de selección Sesgo que ocurre siempre que el investigador selecciona su muestra de una manera sesgada. También conocido como *efecto de selección*.

significancia estadística Una medida de la verosimilitud de que un resultado sea significativo.

significancia práctica En un estudio estadístico, significancia en el sentido que el resultado es asociado con algún curso significativo de acción.

sin correlación Ausencia de cualquier relación evidente entre dos variables.

sujetos En un estudio estadístico, la persona u objeto elegido para la muestra. También vea *participantes*.

tabla de contingencia Vea *tabla de dos entradas*.

tabla de dos entradas Una tabla que muestra la relación entre dos variables listando los valores de una variable en sus renglones y los valores de la otra en sus columnas. También llamada *tabla de contingencia*.

tabla de frecuencias Una tabla que lista todas las categorías de datos en una columna y la frecuencia para cada categoría en otra columna.

tasa de accidentes El número de accidentes debido a alguna causa particular, expresado como una fracción de todas las personas en riesgo por la misma causa. Por ejemplo, una tasa de accidentes de “5 por 1 000 personas” significa que un promedio de 5 entre 1 000 personas sufren un accidente de esta causa particular.

tasa de mortalidad El número de muertes debidas a alguna causa en particular, expresada como una fracción de todas las personas en riesgo para la misma causa. Por ejemplo, una tasa de mortalidad de “5 por 1 000 personas” significa que en promedio 5 de cada 1 000 personas mueren a consecuencia de esta causa particular.

teorema del límite central El teorema establece que para muestras aleatorias (todas del mismo tamaño) de una variable con cualquier distribución (no necesariamente una distribución normal), la distribución de las medias de las muestras, conforme el tamaño de la muestra aumenta, tenderá a ser aproximadamente una distribución normal.

tratamiento Algo dado o aplicado a los miembros del grupo de tratamiento en un experimento.

valor de comparación Un número que se compara con un valor de referencia en el cálculo de una diferencia relativa.

valor de referencia El número que es usado como base para una comparación.

valor esperado El valor medio de los resultados para alguna variable aleatoria.

valor *P* En una prueba de hipótesis, la probabilidad de seleccionar una muestra al menos tan extrema como la observada, suponiendo que la hipótesis nula es verdadera.

valores raros En una distribución de datos, valores que no son probables que ocurran por el azar, como aquellos valores que están a más de dos desviaciones estándar alejadas de la media.

variable Cualquier elemento o cantidad que pueda variar o tomar valores diferentes.

variable de interés En un estudio estadístico, los elementos o cantidades que el estudio busca medir.

variación Cómo están dispersos los datos respecto al centro de una distribución. También vea *desviación estándar*; *rango*; *resumen de cinco números*.

Respuestas

SECCIÓN 1.1

1. Una población es el conjunto completo de personas o cosas que serán estudiadas, mientras que una muestra es un subconjunto de la población. La diferencia radica en que la muestra es sólo parte de la población completa.
3. Una estadística muestral es una característica de una *muestra* encontrada mediante la consolidación o resumen de datos sin procesar. Un parámetro poblacional es una característica de una *población*. Puesto que por lo regular no es práctico medir directamente parámetros poblacionales para poblaciones grandes, inferimos parámetros poblacionales de las estadísticas muestrales medidas.
5. No tiene sentido. 7. No tiene sentido.
9. No tiene sentido.
11. *Muestra*: los 1 002 adultos seleccionados. *Población*: todos los adultos estadounidenses. *Estadística muestral*: 48%. El valor del parámetro poblacional no es conocido, pero es el porcentaje de todos los adultos en Estados Unidos que dijeron, si fueron cuestionados, que estaban a favor.
13. *Muestra*: las distancias de las estrellas seleccionadas. *Población*: las distancias para todas las estrellas en la galaxia. *Estadística muestral*: la distancia media de las estrellas seleccionadas. El valor del parámetro poblacional no es conocido, pero es la distancia media que se obtendría si todas las estrellas de la galaxia fuesen usadas.
15. 45% a 51%. 17. 22.6 a 29.4.
19. Aunque no hay garantía, los resultados sugieren la probabilidad de que gane el republicano por una sólida mayoría, ya que obtendrá entre 55% y 61% de los votos.
21. Con base en la encuesta, se espera que el porcentaje real de votantes esté entre 67% y 73%, que no incluye el valor 61% basado en resultados de votantes reales. Si la encuesta fuese bien realizada, es poco probable que sus resultados fuesen tan diferentes de los resultados reales, implicando que los encuestados mintieron de manera intencional para, al parecer, favorecer a los encuestadores o que sus memorias eran imprecisas.
23. **a. Objetivo**: determinar cómo ejecutivos satisfechos están con sus elecciones de carrera. *Población*: el conjunto completo de todos los ejecutivos. *Parámetro poblacional*: el porcentaje de todos los ejecutivos que dirían, si se les pregunta, que si pudiesen iniciar otra vez su carrera elegirían un campo diferente. **b. Muestra**: los 1 733 ejecutivos seleccionados para la encuesta. *Datos sin procesar*: respuestas individuales a la pregunta. *Estadístico muestral*: 51%. **c.** 48% a 54%
25. **a. Objetivo**: determinar el porcentaje de adultos que están desempleados. *Población*: el conjunto completo de todos los adultos. *Parámetro poblacional*: el porcentaje de todos los adultos que están desempleados. **b. Muestra**: la muestra de adultos seleccionados para la encuesta. *Datos sin procesar*: respuestas individuales a la pregunta. *Estadística muestral*: 4.6% **c.** 4.4% a 4.8%
27. *Paso 1*: **Objetivo**: identificar el porcentaje de todos los conductores con licencia que usaron un teléfono celular al menos una vez mientras conducían durante la semana pasada. *Paso 2*: elegir una muestra de conductores con licencia. *Paso 3*: encuestar a una muestra de conductores con licencia para determinar cuántos de ellos usaron un teléfono celular al menos una vez mientras estaban manejando durante la semana pasada. Para esta muestra, encontrar el porcentaje de conductores con licencia que utilizaron un teléfono celular al menos una vez mientras conducían durante la semana pasada. *Paso 4*: utilizar las técnicas de la ciencia estadística para hacer una inferencia acerca del porcentaje de conductores con licencia que usaron un teléfono celular al menos una vez mientras conducían la semana pasada. *Paso 5*: con base en el valor probable del parámetro poblacional, formar una conclusión acerca del porcentaje de conductores con licencia que usaron un teléfono celular al menos una vez durante la semana pasada.
29. *Paso 1*: **Objetivo**: identificar el peso medio de todos los pasajeros de aerolíneas comerciales. *Paso 2*: elegir una muestra de pasajeros de aerolíneas. *Paso 3*: pesar a cada pasajero seleccionado, y luego determinar la media de estos pesos. *Paso 4*: utilice las técnicas de la ciencia estadística para hacer una inferencia acerca del peso medio de todos los pasajeros de aerolíneas. *Paso 5*: con base en el valor probable de la media poblacional, forme una conclusión acerca del peso medio de todos los pasajeros de aerolíneas.

SECCIÓN 1.2

1. Un censo es la colección de datos de cada miembro de la población; una muestra es una colección de datos de sólo parte de la población.
3. Con un muestreo por conglomerados, seleccionamos a todos los miembros de subgrupos (o conglomerados) seleccionados aleatoriamente; con muestreo estratificado, seleccionamos muestras de cada uno de los diferentes subgrupos (o estratos).
5. No tiene sentido. 7. Tiene sentido.

9. Un censo es práctico. La población consiste en el número reducido de jugadores del equipo de Los Ángeles Lakers y es fácil obtener sus estaturas.
11. Un censo no es práctico. El número de instructores de estadística es grande y sería muy difícil aplicarles a todos un examen de CI.
13. La muestra son los tiempos de servicio de los cuatro senadores seleccionados. La población son los tiempos de servicio de los 100 senadores. Éste es un ejemplo de muestreo aleatorio (o muestreo aleatorio simple). Sin embargo, puesto que la muestra es tan pequeña, no es probable que sea representativa de la población.
15. La muestra son los 1012 adultos seleccionados aleatoriamente. La población es el conjunto completo de todos los adultos estadounidenses. Puesto que la muestra es muy grande y fue obtenida por una compañía respetable, es probable que la muestra sea representativa de la población.
17. La muestra *c* es la más representativa, ya que es probable que la lista represente a la mayoría de la gente y no haya razón para pensar que la gente con los primeros 1000 números diferiría en alguna forma de otras personas. La muestra *a* está sesgada porque incluye sólo propietarios de vehículos caros. La muestra *b* está sesgada porque sólo incluye personas de una región geográfica. La muestra *d* está sesgada porque incluye una muestra autoseleccionada que consiste de personas que es más probable que tengan fuertes sentimientos acerca del tema de deuda de tarjetas crédito
19. Puesto que la crítica de cine trabaja indirectamente con Disney, podría estar más inclinada a emitir una revisión favorable.
21. Los científicos universitarios reciben pago de Monsanto, así que podrían estar inclinados a complacer a la compañía con el deseo de obtener proyectos en el futuro. Así, podría haber una inclinación a proporcionar resultados favorables.
23. La muestra es aleatoria simple y es probable que sea representativa ya que no hay sesgo en el proceso de selección.
25. El muestreo es por conglomerados y es probable que sea representativo, aunque el método exacto de selección de las estaciones para encuestar podría afectar por sesgo a la muestra.
27. La muestra es por conveniencia. Es probable que esté sesgada porque la muestra consiste de miembros de la familia que quizá tengan características físicas similares y no es probable que esas características sean representativas de la población general.
29. La muestra es estratificada. Es probable que esté sesgada porque la gente de esos grupos de edad no está distribuida uniformemente en toda la población. Sin embargo, podría ajustarse para reflejar la distribución de edades de la población.
31. La muestra es sistemática. Es probable que sea representativa porque no hay nada acerca del orden alfabético que posibilite tener como resultado una muestra sesgada.
33. La muestra es estratificada. Es probable que esté sesgada porque la población no tiene igual número de personas en cada una de las tres categorías. Sin embargo, los resultados podrían ponderarse para reflejar la distribución real de la población.
35. La muestra es por conveniencia. Es probable que esté sesgada porque es una muestra por autoselección y consiste de aquéllos con fuerte sentimiento acerca del tema.
37. La muestra es aleatoria simple, por lo que es probable que sea representativa.
39. Muestreo aleatorio simple del grupo estudiantil, con un tamaño de muestra suficientemente grande para obtener resultados significativos.
41. El muestreo por conglomerados de registros de muertes debe dar buena información. Los conglomerados podrían estar basados en género, ubicación geográfica o ambas.

SECCIÓN 1.3

1. Un placebo es físicamente semejante a un tratamiento, pero carece de ingredientes activos, de modo que no debe producir efecto alguno. Un placebo es importante de modo que los resultados de los sujetos a quienes se les da un tratamiento real puedan compararse con los resultados de los sujetos a quienes se les dio un placebo.
3. La confusión es la mezcla de efectos de diferentes factores, de modo que efectos específicos no pueden atribuirse a factores específicos. Por ejemplo, si a hombres se les da el tratamiento y a mujeres se les dan placebos, no puede determinarse si los efectos se deben al tratamiento o al sexo del participante.
5. Tiene sentido usar un experimento doble ciego, pero el diseño experimental es delicado ya que los sujetos claramente pueden ver la ropa que están usando y los evaluadores también ven esos colores. El experimento podría hacerse ciego no diciendo a los sujetos acerca del experimento, de modo que el conocimiento del color de su ropa no afectaría sus resultados. Puesto que es difícil hacer ciego el experimento para los evaluadores, es importante obtener resultados con base en medidas objetivas para el juicio de los evaluadores.
7. El efecto del experimentador ocurre cuando un investigador o experimentador influye de alguna manera a los sujetos por medio de factores tales como expresión facial, tono de voz o actitud. El efecto del experimentador puede evitarse usando un experimento doble ciego o ciego sencillo, de modo que quienes evalúan los resultados no sepan a cuáles sujetos se les dio el tratamiento y a cuáles se les dio el placebo.

9. Estudio observacional ya que los sujetos fueron probados, pero no se les dio ningún tratamiento. La variable de interés representa el resultado *correcto* o *incorrecto* para cada ensayo.
11. Experimento, ya que el grupo de tratamiento consiste de los sujetos a quienes se les dieron los brazaletes magnéticos y el grupo de control consiste en los sujetos a quienes se les dio el brazalet que no tiene magnetismo. Es probable que no se pueda usar experimento ciego con este estudio, ya que los pasajeros con facilidad podrían probar los brazaletes para ver si son magnéticos (colocándolos cerca de metal). La variable de interés es si el pasajero experimenta malestar por el movimiento.
13. Estudio observacional retrospectivo, que examina cómo una característica determinada antes de nacer (mellizos idénticos o fraternos) afecta las capacidades mentales posteriores. La variable de interés es la diferencia entre el nivel de medida de las capacidades mentales en las dos personas que son mellizos.
15. Meta-análisis. La variable de interés es si el sujeto desarrolló cáncer de próstata.
17. Experimento. El grupo de tratamiento consiste en el maíz modificado genéticamente; el grupo de control consiste en el maíz que no está modificado genéticamente. La variable de interés es la cantidad medida de insecticida que es liberada a través de las raíces.
19. Experimento. El grupo de tratamiento consiste de aquéllos tratados con imanes; el grupo de control consiste de aquéllos a los cuales se les dieron dispositivos no magnéticos. La variable de interés es el nivel de medida de dolor de espalda.
21. Es probable que ocurra confusión. Si hay diferencia en los efectos de los dos grupos, no hay manera de saber si esas diferencias son atribuibles al tratamiento (fertilizantes o irrigación) o al tipo de región (húmeda o seca). Esta confusión podría evitarse usando bloques de árboles fertilizados en ambas regiones, y usando bloques de árboles irrigados en ambas regiones.
23. Es probable la confusión. Si hay diferencias en las cantidades de gasolina consumida no hay manera de saber si estas diferencias son debidas al octanaje de la gasolina o al tipo de vehículo. La confusión puede evitarse usando gasolina de 87 octanos en la mitad de camionetas y la mitad de vehículos deportivos y usando gasolina de 91 octanos en los otros vehículos.
25. Es posible la confusión por un efecto placebo y/o efecto del experimentador. Puede evitarse con un experimento doble ciego.
27. Es posible la confusión por un efecto placebo y/o efecto del experimentador, con las pelotas de tenis desempeñando el papel de placebo. Debe ser mejor usar las pesas y las pelotas de tenis con los mismos sujetos en tiempos diferentes, con el orden mezclado.
29. El grupo de control consiste en aquellos que no escucharon Beethoven y el grupo de tratamiento consiste en aquellos que escucharon Beethoven. Podría usarse un experimento ciego mediante la codificación de sujetos, de modo que quienes midan la inteligencia no estén influidos por su conocimiento acerca de los participantes.
31. El grupo de control consiste en automóviles que usan gasolina sin el aditivo etanol y el grupo de tratamiento consiste de automóviles que usan gasolina con el aditivo etanol. No es necesario que sea ciego para los automóviles y quizá no sea necesario para los investigadores, ya que el millaje probablemente sea medido con herramientas objetivas. Además, los mismos automóviles podrían usarse con y sin el aditivo por lo que las diferencias extrañas serían eliminadas.

SECCIÓN 1.4

1. Revisión de pares es un proceso en el que expertos en un campo evalúan un reporte de investigación antes que éste sea publicado. Es útil para dar credibilidad a la investigación porque implica que otros expertos coinciden en que la investigación fue realizada de manera apropiada.
3. Cuando los participantes se seleccionan ellos mismos para una encuesta, aquéllos con opiniones sólidas acerca del tema que será encuestado es más probable que participen, y por lo común es un grupo que no es representativo de la población general.
5. No tiene sentido. 7. No tiene sentido.
9. Puesto que los investigadores son del departamento de relaciones públicas, pueden estar sesgados, por lo que la directriz 2 es la más relevante.
11. Puesto que “buena” no está bien definida porque quizá sea difícil medir “ética buena”, la directriz 4 es importante.
13. La muestra es autoseleccionada, por lo que el sesgo de participación es un tema grave. La directriz 3 es la más relevante.
15. La redacción de la pregunta está sesgada y tiende a obtener respuestas negativas, por lo que la directriz 6 es la más relevante.
17. Puesto que gran parte de los fondos fueron proporcionados por Mars y la Asociación de Fabricantes de Chocolate, los investigadores están más inclinados a dar resultados favorables. El sesgo podría evitarse si los investigadores no son pagados por los fabricantes de chocolates. Si esa es la única forma en que la investigación podría hacerse, los investigadores deberían haber instituido procedimientos para asegurar que todos los resultados, incluso los negativos, fuesen publicados.
19. Puesto que los encuestados son autoseleccionados, probablemente representen a aquellos que tienen fuertes sentimientos acerca del tema. Un mejor método de

muestreo, como el aleatorio simple usado para la mayoría de las encuestas, es necesario.

21. La frase “está mal” en la primera pregunta podría ser engañosa. Algunas personas podrían creer que el aborto está mal pero aún estar a favor de la elección. La segunda pregunta también podría ser confusa, cuando algunas personas podrían pensar *que con la ayuda de su doctora* significa que la vida de la mujer está en peligro, lo que podría alterar la opinión acerca del aborto en esta situación. Grupos que se oponen al aborto sería probable que citasen los resultados de la primera pregunta, mientras que grupos que favorecen la elección sería probable que citasen los resultados de la segunda pregunta.
23. La primera pregunta requiere un estudio de citas por internet. La segunda pregunta incluye un estudio de personas casadas para determinar si su primera cita fue por internet. El segundo grupo es mucho más limitado que el primero.
25. La primera pregunta incluye un estudio de estudiantes universitarios en general. La segunda pregunta incluye un estudio de quienes se emborrachan. La primera pregunta podría dirigirse mediante una encuesta a estudiantes universitarios. La segunda pregunta sería dirigida a encuestar tomadores que se emborrachan, un grupo que sería mucho más difícil de encuestar.
27. El encabezado se refiere a drogas, mientras que la historia se refiere a uso de drogas, beber o fumar. Puesto que *drogas* son consideradas, por lo general, como drogas diferentes del cigarro o del alcohol, el encabezado es muy engañoso.
29. No se da información alguna acerca del significado de *confianza*. El tamaño de la muestra y el margen de error no son proporcionados.
31. No se da información para justificar el enunciado que “más” compañías tratan de apostar al pronóstico del clima. Si sólo las cuatro compañías citadas son nuevas, el aumento es relativamente insignificante.

EJERCICIOS DE REPASO DEL CAPÍTULO 1

1. **a.** 35% a 41% **b.** Todos los adultos en Estados Unidos **c.** Estudio observacional porque los sujetos no fueron tratados o modificados de manera alguna. La variable de interés es *armas en casa*, que para este estudio puede tomar dos valores: *sí* o *no*. **d.** El valor es un estadístico muestral porque está basado en la muestra de 1012 adultos, no la población de todos los adultos. **e.** No, porque sería una muestra autoseleccionada con una participación quizá sesgada. **f.** Existen muchos procedimientos diferentes para seleccionar una muestra aleatoria simple de adultos en Estados Unidos, pero cualquier procedimiento debe diseñarse para asegurar que todas las muestras de 1012 adultos tengan la misma oportunidad de ser elegidas. Un

procedimiento sencillo es compilar una lista numerada de todos los adultos en Estados Unidos, usar una computadora para generar de manera aleatoria 1012 números diferentes, y luego seleccionar los sujetos correspondientes a los números seleccionados. **g.** Seleccionar una muestra de hogares en cada estado. **h.** Seleccionar todos los hogares en varios distritos electorales elegidos aleatoriamente. **i.** Seleccionar cada décimo hogar, por dirección, en cada calle en una ciudad. **j.** Seleccionar los hogares de sus compañeros de clase.

2. **a.** Una muestra aleatoria simple es una muestra elegida de manera tal que toda muestra del mismo tamaño tenga la misma oportunidad de ser elegida. **b.** No, ya que no toda muestra de tamaño 2007 de personas tiene la misma oportunidad de ser elegida. Por ejemplo, es imposible seleccionar la muestra con 2007 personas en la misma unidad primaria de muestreo. **c.** Repita el proceso de elegir aleatoriamente una unidad primaria de muestreo y luego seleccione uno de sus miembros. Si alguien es seleccionado más de una vez, ignore las selecciones adicionales.
3. **a.** No, ya que no hay información acerca de la ocurrencia de dolor de cabeza entre gente que no usó Zocor. **b.** Puesto que la tasa de dolor de cabeza es menor entre usuarios de Zocor, parece que los dolores de cabeza no son una reacción adversa del uso de Zocor. **c.** Con prueba ciega, los participantes del ensayo no saben si ellos tienen Zocor o un placebo y aquellos que evalúan los resultados tampoco lo saben. **d.** Experimento, porque a los sujetos se les da un tratamiento. **e.** Ocurre un efecto del experimentador, si éste de alguna manera influye en los sujetos a través de factores tales como expresión facial, tono de voz o actitud. Puede evitarse por medio de prueba ciega.
4. **a.** La segunda pregunta, ya que la palabra *bienestar* tiene connotaciones negativas. **b.** La primera pregunta, ya que es más probable obtener respuestas negativas. **c.** Eso es un juicio muy subjetivo. Algunos encuestadores profesionales se oponen a tales preguntas que deliberadamente están sesgadas, pero otros consideran que tales preguntas pueden usarse. Una consideración importante es que las preguntas de la encuesta pueden modificar cómo piensa la gente, y tal modificación no debe ocurrir sin conciencia o acuerdo.

CUESTIONARIO DEL CAPÍTULO 1

1. a 2. c 3. a 4. b 5. c 6. a 7. c
8. c 9. b 10. b 11. b 12. b 13. b
14. c 15. b

SECCIÓN 2.1

1. Los datos cualitativos consisten en valores que pueden colocarse en categorías no numéricas diferentes, mientras

que los datos cuantitativos consisten en valores que representan conteos o medidas.

3. Sí. Los datos consisten de cualidades o cantidades (números), por lo que todos los datos son cualitativos o cuantitativos.
5. Los grupos sanguíneos son cualitativos porque ellos no miden o cuentan algo.
7. Las estaturas son cuantitativas ya que consisten en mediciones.
9. Las duraciones de películas son cuantitativas ya que consisten en mediciones.
11. Los programas de televisión son cualitativos ya que no miden ni cuentan algo.
13. El número es cuantitativo ya que consiste de un conteo.
15. Los salarios son cuantitativos ya que consisten en conteo de dinero.
17. Los números son discretos porque sólo se usan los números de conteo y ningún valor entre ellos es posible.
19. Los números son discretos ya que son conteos. Sólo los números de conteo son usados y ningún valor entre ellos es posible.
21. Los tiempos son datos continuos ya que pueden tomar cualquier valor entre algún rango de valores.
23. Las calificaciones de exámenes son discretas ya que sólo pueden ser números de conteo.
25. Las velocidades son datos continuos ya que pueden tomar cualquier valor dentro de algún rango de valores.
27. Los números son discretos ya que sólo pueden ser números de conteo.
29. Razón 31. Nominal 33. Ordinal 35. Ordinal
37. Nominal 39. Razón
41. El nivel de razón no aplica. La razón no tiene significado porque las estrellas no miden ni cuentan algo. Las diferencias entre valores de estrellas no tienen significado.
43. El nivel de razón aplica. La velocidad de 450 mi/h es tres veces la velocidad de 150 mi/h.
45. El nivel de razón aplica. La razón de “dos veces” tiene sentido.
47. El nivel de razón aplica. El monto del salario \$150 000 es el doble del salario de \$75 000, por lo que la razón de “dos veces” tiene sentido.
49. Los datos son cuantitativos y están al nivel de razón de medida. Los datos son continuos. Los tiempos tienen un punto cero inicial natural y pueden ser cualquier valor dentro de un rango particular.

51. Los datos son cualitativos y están en el nivel nominal de medida. Los números son diferentes maneras de expresar los nombres y no miden ni cuentan algo.
53. Los datos son cuantitativos y están al nivel de intervalo de medida. Los datos son discretos ya que consisten sólo de números enteros. Los años se miden desde una referencia arbitraria (el año 0), no un punto inicial natural. Las diferencias entre años tienen sentido, pero las razones no.
55. Los datos son cualitativos y están en el nivel ordinal de medida. Las clasificaciones consisten en un orden, pero no representan cuentas o medidas.

SECCIÓN 2.2

1. Puesto que es el resultado de un problema registrado en el sistema de medida, es un error sistemático. No es un error aleatorio porque no es el resultado del azar ni de eventos inherentemente impredecibles en el proceso de medición.
3. Puesto que la estatura registrada tiene tantos decimales, es muy precisa. Puesto que la estatura registrada no es muy cercana a la estatura verdadera, no es muy exacta.
5. No tiene sentido. 7. Tiene sentido
9. Los errores tienden a resultar de errores aleatorios, pero la deshonestidad tiende a dar resultados sistemáticos que benefician al contribuyente.
11. Puesto que alrededor de la mitad de los tornillos son más largos que 25 mm y la mitad son más cortos que 25 mm, esto parece ser un error aleatorio.
13. Los errores aleatorios ocurren cuando ingresos reportados se registran de manera incorrecta o cuando los encuestados no conocen sus ingresos exactos. Los errores sistemáticos ocurren cuando las personas reportan ingresos más altos de modo que parezca que son más exitosos.
15. Los errores aleatorios ocurren con escalas inconsistentes o errores en la lectura de las escalas o al registrar los datos. Los errores sistemáticos ocurren con una escala que de manera consistente proporciona lecturas mayores o menores.
17. Los errores aleatorios ocurren con una pistola de radar inconsistente o con errores honestos por el oficial que registra las velocidades. Los errores sistemáticos ocurren con una pistola de radar que está mal calibrada de modo que es consistente en dar lecturas mayores o menores.
19. Los errores aleatorios ocurren con errores en los cálculos. Los errores sistemáticos ocurren por el reporte faltante de cajetillas de cigarros que se obtienen de manera ilegal sin estampillas de impuestos.
21. *Error absoluto:* \$1 750. *Error relativo:* 141% (usando el error de \$1 750 y el monto correcto de la factura de \$1 245).
23. *Error absoluto:* -2 mi/gal. *Error relativo:* -8.3%.

- 25. a.** Estos errores son aleatorios. **b.** El promedio es la mejor elección para minimizar los errores aleatorios. **c.** Errores sistemáticos podrían surgir, por ejemplo, de un problema con el dispositivo de medición o un problema en la definición de la “longitud” de una habitación. **d.** Medidas promedio no reducirán los errores sistemáticos.
- 27.** La báscula del Departamento de Transporte es más precisa ya que su peso de 3 298.2 lb es más detallado que el otro peso de 3 250 lb. La báscula del fabricante es más exacta ya que su peso de 3 250 lb es más cercano al valor verdadero. (El peso de 3 250 lb tiene un error de 23 lb, pero el peso de 3 298.2 lb tiene un error de 25.2 lb).
- 29.** La báscula digital del gimnasio es más precisa y más exacta que la báscula de la clínica (suponiendo que su peso real es el que usted cree).
- 31.** El número dado es muy preciso, pero probablemente no sea exacto. En la actualidad no podemos medir la población con esa precisión y la incertidumbre en 1860 era mayor.
- 33.** El número dado es muy preciso pero probablemente no es muy exacto. Hay mucha gente en China que no es contada. Es probable que el censo de cualquier país esté equivocado por una cantidad considerable, a consecuencia de las dificultades inherentes en la realización de un censo nacional. La población de China probablemente cambia durante el transcurso del año.
- 35.** Es fácil medir con precisión la altura de una estructura con un grado razonable de precisión, tal como al 1/10 de pie más cercano, pero el número dado tiene demasiada precisión (implica el conocimiento de la altura a ¡menos de un átomo!), así que la afirmación no es creíble.
- 37.** El número de estudiantes universitarios cambia de manera constante con nuevas inscripciones y bajas, por lo que el número dado debe ser una estimación. El número dado sugiere que es preciso sólo al millón más próximo, lo cual parece muy posible para una buena estimación. Así que es creíble, aunque necesitaríamos más información para saber si realmente es exacto.

SECCIÓN 2.3

- 1.** Para alumnos de décimo grado, la nueva tasa de fumadores es 18.3% y el porcentaje de aumento es 45%. Por tanto, la tasa anterior por 1.45 es igual a la nueva tasa de 18.3%, por lo que la tasa anterior de alumnos de 10o. grado es 12.6%. Para alumnos de 8o. grado, la nueva tasa de fumadores es 10.4% y el porcentaje de aumento es 44%. Por tanto, la tasa anterior por 1.44 es igual a la nueva tasa de 10.4%, así que la tasa anterior para alumnos de 8o. es 7.2%.
- 3.** El enunciado implica de manera incorrecta que el error puede ser subir 1.2% de 5%. Puesto que 1.2% de 5% es

0.06%, el enunciado implica de manera incorrecta que el error puede subir 0.06% (de 4.94% a 5.06%); el error real puede ser 1.2 puntos porcentuales alejado de 5% (de 3.8% a 6.2%).

- 5.** No tiene sentido. **7.** Tiene sentido.
- 9. a.** 2/5, 0.4; 40% **b.** 150/100, o 3/2; 1.50; 150%
c. 1/4; 0.25; 25% **d.** 30/100, o 3/10; 0.30; 30%
- 11. a.** 93% **b.** 44% **c.** 41% **d.** 81%
- 13.** -35% (Hay un 35% de decrecimiento desde 1990).
- 15.** 115% (Hay un incremento de 115% desde 1980).
- 17.** 61%. El *Wall Street Journal* tiene 61% más circulación que el *New York Times*.
- 19.** -19%. O'Hare dio servicio a 19% menos pasajeros que Hartsfield.
- 21.** 66
- 23.** 834
- 25.** 140%. El camión pesa 100% del peso del automóvil más otro 40%.
- 27.** 80%. La población de Montana es 100% de la población de New Hampshire menos 20% de la población de New Hampshire.
- 29.** Sí. Tres puntos porcentuales corresponden a un rango desde 86% a 92%, que era la intención. Un margen de error de 3% correspondería a 3% de 89%, que es $0.03 \times 0.89 = 0.0267$, pero no es lo que se quería decir.
- 31.** -15.5 puntos porcentuales; -22.7%
- 33.** 22 puntos porcentuales; 56.4%

SECCIÓN 2.4

- 1.** Un número índice es una razón sin unidades. El número aparece como un costo real de gasolina en 2007, no un número índice.
- 3.** Sí. El índice de precios al consumidor está basado en los precios de bienes, servicios y vivienda, por lo que aumentos en esos precios resultarán en un aumento en el índice de precios al consumidor (IPC).
- 5.** 573.2 **7.** \$1.12
- 9.** 93.3, 100, 181.7, 383.3, 386.2, 496.8, 740.4
- 11.** \$21.69
- 13.** El costo en 2004 es 466.4% del costo en 1980, pero el IPC en 2004 es 229.2% del IPC de 1980, por lo que el costo de la universidad se elevó mucho más que los costos de bienes, servicios y vivienda comunes.
- 15.** El precio de las casas en 2004 es 206.5% de la cantidad en 1990, pero el IPC en 2004 es 144.5% del IPC de 1990, por

lo que el costo de casas se elevó a una tasa inferior a los costos de bienes, servicios y vivienda comunes.

17. \$582 000; \$180 000 19. \$1 591 667; \$1 491 667

EJERCICIOS DE REPASO DEL CAPÍTULO 2

1. a. 706 b. Discreta, una persona no puede sobrevivir de manera fraccionaria. c. 23.89% d. 64
e. Razón f. Nominal
2. a. 860 b. 37% c. Ordinal, puesto que hay un orden. d. Puesto que la encuesta utiliza encuestados que ellos mismos deciden participar, la muestra es de autoselección y es probable que no refleje con precisión la opinión de la población.
3. El gasto en salud en 2004 es 2 150% del monto de 1973, pero el IPC en 2004 es 325.5% del IPC de 1973, por lo que el gasto en salud creció a una tasa mucho mayor que la tasa general para bienes, servicios y vivienda.
4. a. \$2.78 b. \$5.77 c. El aumento del salario mínimo no se mantuvo con la inflación, por lo que el salario mínimo real de 2006 de \$5.15 tiene el mismo poder de compra que \$4.14 en 1996. Los trabajadores que ganaban el salario mínimo de \$4.75 en 1996 tenían mayores ingresos relativos que los trabajadores que ganaban \$5.15 en 2006.

CUESTIONARIO DEL CAPÍTULO 2

1. Nominal 2. Continuos 3. Razón
4. -10 cm 5. -18.1% 6. 52% 7. 52
8. 178 cm, 179.18 cm 9. 123.7 10. \$52 972

SECCIÓN 3.1

1. Una tabla de frecuencias tiene dos columnas, una para las categorías y otra para las frecuencias. Las categorías son los diferentes valores que una variable puede tener (por ejemplo, diferente color de ojos o sabores de un helado). Las frecuencias son los números de datos (total) en cada categoría.
3. 2, 11, 25, 37, 40 5. No tiene sentido.
7. No tiene sentido.

9.

Calificación	Frecuencia	Frecuencia relativa	Frecuencia acumulada
A	4	16.7%	4
B	7	29.2%	11
C	8	33.3%	19
D	3	12.5%	22
E	2	8.3%	24
Total	24	1 = 100%	24

11.

Peso (libras)	Frecuencia	Frecuencia relativa	Frecuencia acumulada
0.7900-0.7949	1	1/36	1
0.7950-0.7999	0	0	1
0.8000-0.8049	1	1/36	2
0.8050-0.8099	3	3/36	5
0.8100-0.8149	4	4/36	9
0.8150-0.8199	17	17/36	26
0.8200-0.8249	6	6/36	32
0.8250-0.8299	4	4/36	36
Total	36	100%	36

13.

Edad	Núm. de actores
20-29	0
30-39	12
40-49	13
50-59	5
60-69	3
70-79	1

15.

Categoría	Frecuencia	Frecuencia relativa
A	13	24%
B	9	18%
C	12	24%
D	11	22%
F	6	12%
Total	50	100%

17. a. 200 b. 142 c. 16% d. 0.135, 0.155, 0.210, 0.200, 0.140, 0.160 e. 27, 58, 100, 140, 168, 200

19. a.

Calificación	Frecuencia	Frecuencia relativa
0-2	20	38.5%
3-5	14	26.9%
6-8	15	28.8%
9-11	2	3.8%
12-14	1	1.9%
Total	52	100%

b.

Calificación	Frecuencia	Frecuencia relativa
0-2	33	63.5%
3-5	19	36.5%
6-8	0	0%
9-11	0	0%
12-14	0	0%
Total	52	100%

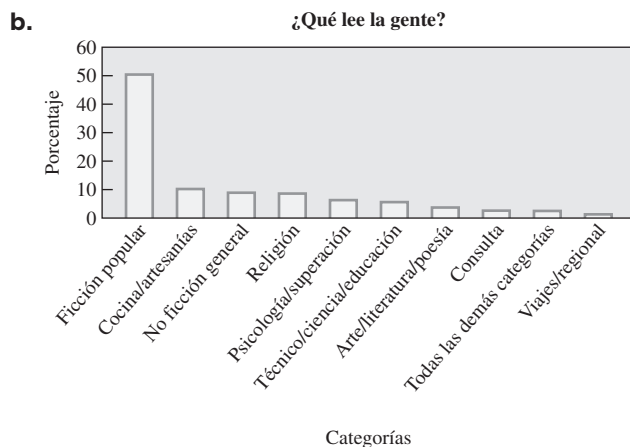
- c. El teclado Dvorak parece ser más eficiente porque tiene calificaciones más bajas y menos altas.

SECCIÓN 3.2

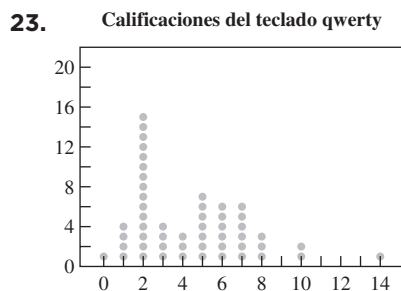
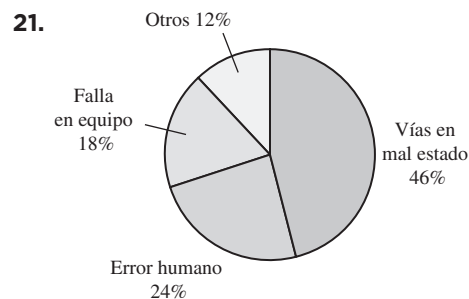
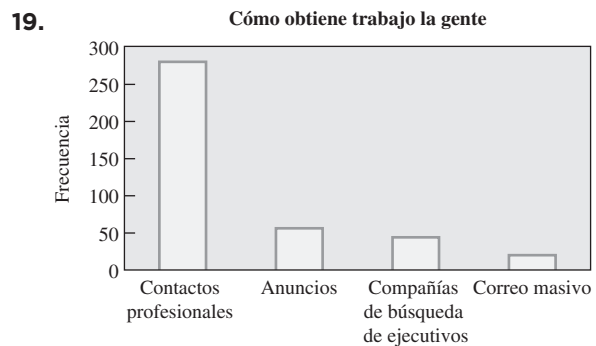
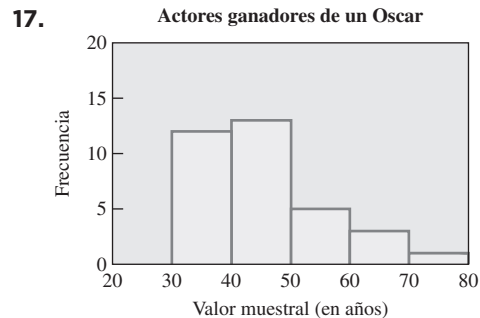
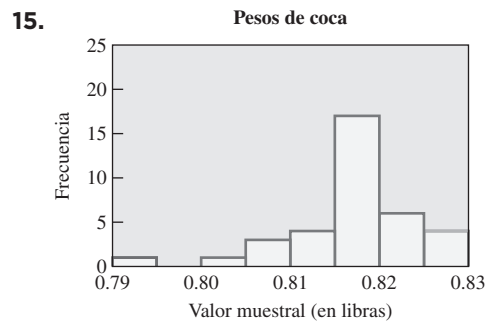
1. La distribución de datos es la forma en que los valores están esparcidos sobre todos los valores posibles.
El histograma proporciona una gráfica que tiene una forma, y es mucho más sencillo entender la forma de la distribución a partir de la gráfica que de una lista de valores.
3. Los datos deben ordenarse en alguna sucesión temporal, y la gráfica de serie de tiempo ayuda a revelar tendencias o patrones en el tiempo.
5. No tiene sentido. 7. Tiene sentido.
9. Un histograma funcionaría bien para mostrar las frecuencias de las diferentes categorías de calificaciones.
11. Un diagrama de Pareto o diagrama circular funcionarían bien, pero el diagrama de Pareto hace un mejor trabajo para mostrar las causas más comunes de muerte.

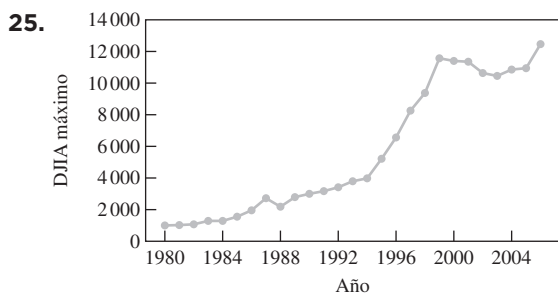
13. a.

Categoría de lectura	Frecuencia relativa
Ficción popular	50.4%
Cocina/artesanías	10.2%
No ficción general	8.9%
Religión	8.6%
Psicología/superación	6.3%
Técnico/ciencia/educación	5.6%
Arte/literatura/poesía	3.7%
Consulta	2.6%
Todas las demás categorías	2.5%
Viajes/regional	1.3%



- c. El diagrama de Pareto hace más sencillo ver cuáles categorías son las más populares.





Al paso de los años existe una tendencia de crecimiento de las acciones, por lo que el mercado de valores parece ser una buena inversión.

27.

Tallo	Hojas
6	7
7	25
8	5899
9	09
10	0

Las longitudes de los renglones son similares a las alturas de las barras en un histograma; renglones más largos de datos corresponden a frecuencias mayores.

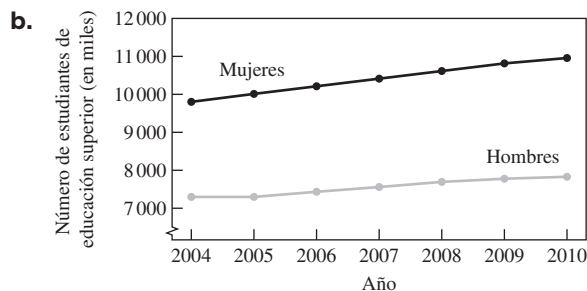
SECCIÓN 3.3

1. El uso de barriles de petróleo no es apropiado porque los datos no son tridimensionales. Sería mejor ilustrar los datos con una gráfica simple de barras.

3. Los datos geográficos son datos sin procesar que corresponden a diferentes ubicaciones geográficas. Dos ejemplos de exhibición de información geográfica son mapas coloreados y mapas de contornos.

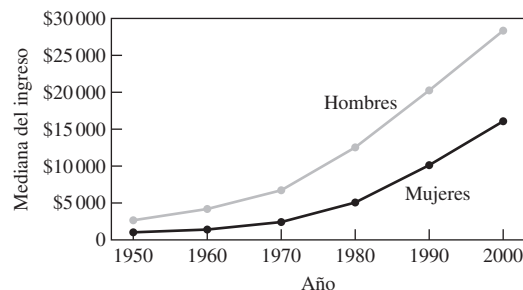
5. No tiene sentido. 7. No tiene sentido.

9. a. Las mujeres exceden de manera consistente el número de hombres. Los números de ambos géneros están creciendo gradualmente con el tiempo.



11. a. La mediana de los ingresos de los hombres es consistentemente mayor que la de las mujeres, y tanto hombres como mujeres han aumentado sus ingresos al paso del tiempo. Al comparar las alturas de las barras de izquierda a derecha, vemos que las razones de ingresos de hombres al ingreso de mujeres parece que están

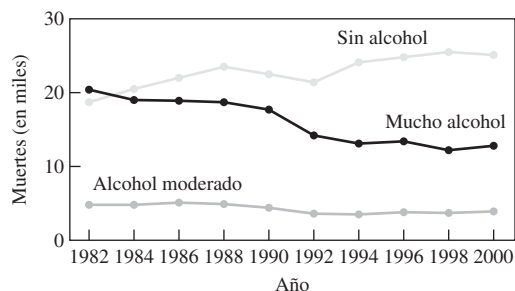
disminuyendo. El par de más a la izquierda muestra que el ingreso de los hombres es el doble del ingreso de las mujeres, pero el par de más a la derecha parece mostrar que el ingreso de hombres no es tan grande como el doble del ingreso de las mujeres. b. La gráfica de líneas hace más sencillo examinar la tendencia con el tiempo. Además, la gráfica de línea está menos abarrotada, por lo que la información es más fácil de entender e interpretar.

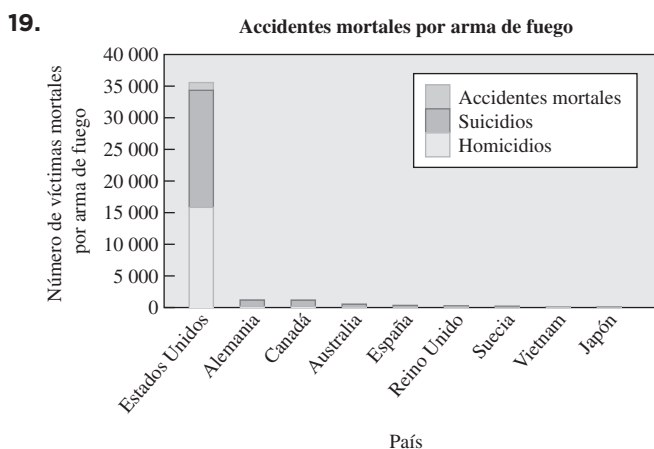


13. Esta gráfica apilada muestra *porcentajes* del total del presupuesto. Por tanto, su altura total siempre es 100% y no podemos determinar directamente la cantidad real de dinero que gasta el gobierno. Sin embargo, es fácil ver la tendencia a largo plazo de cómo el gobierno gasta su dinero. Por ejemplo, la proporción del presupuesto federal que va a la defensa nacional cayó de manera sustancial, de alrededor de 50% del presupuesto federal en 1960 a alrededor de 13% en 2005. Durante el mismo periodo, la proporción del presupuesto que va a pagos de individuos (por ejemplo, seguridad social, seguro médico y bienestar) se elevó drásticamente, de casi 30% a casi 70%. La proporción gastada en interés neto creció más del doble.

15. Parece que existe una tendencia general de tasas de mortalidad más alta por melanoma en estados del sur y estados del oeste de Estados Unidos. Esto podría ser el resultado de que la gente en estos estados pasa más tiempo en el exterior, exponiéndose a la luz solar. Como investigador, usted podría estar interesado particularmente en regiones que se desvían de las tendencias generales. Por ejemplo, un condado al este del estado de Washington destaca con una tasa alta de mortalidad por melanoma. Usted primero podría necesitar verificar que el dato es exacto y no un error de algún tipo. Si es exacto, podría querer determinar por qué este condado tiene una tasa más alta de mortalidad que los condados que lo rodean.

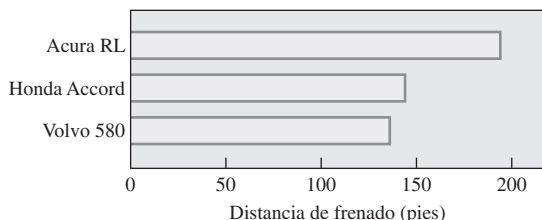
17. Alcohol involucrado en accidentes automovilísticos mortales





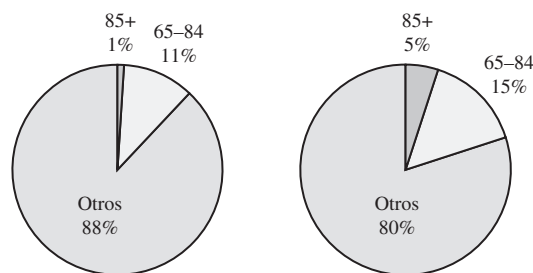
SECCIÓN 3.4

- Recortar la parte inferior de la gráfica distorsiona la percepción del lector, causando que sólo vea parte de la gráfica en lugar de la gráfica completa, lo que da una mejor perspectiva de toda la información. Los cambios tienden a ser exagerados.
- En lugar de usar intervalos que cambian en la misma cantidad, una escala exponencial utiliza intervalos que cambian por potencias de algún número (con frecuencia 10). Por ejemplo, si se usan potencias de 10, intervalos iguales en la escala representan 1, 10, 100, y así sucesivamente. Las escalas exponenciales son útiles para mostrar datos que varían en un rango muy amplio.
- La gráfica es engañosa porque la escala horizontal no inicia en cero. La distancia de frenado para el Acura es 42% mayor que para el Volvo



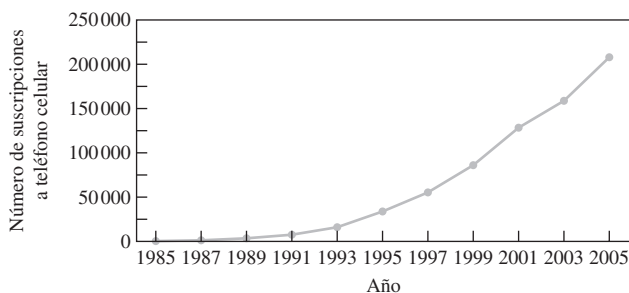
- La cantidad de petróleo usado por cada país parece que está relacionada con el volumen de los barriles en el pictograma, cuando en realidad está relacionada con la altura de los barriles. El consumo en Estados Unidos es alrededor de cuatro veces el de Japón, no 64 veces como sugieren los volúmenes de los barriles.
- a, b.** A consecuencia de la apariencia tridimensional de las gráficas circulares, los tamaños de los sectores en la página no coinciden con los porcentajes. En cambio, muestran cómo los sectores se verían si el círculo completo se inclinase a un ángulo. Esta distorsión hace difícil ver la verdadera relación entre las categorías.

c. Distribución de edades en 1990 Distribución de edades en 2050

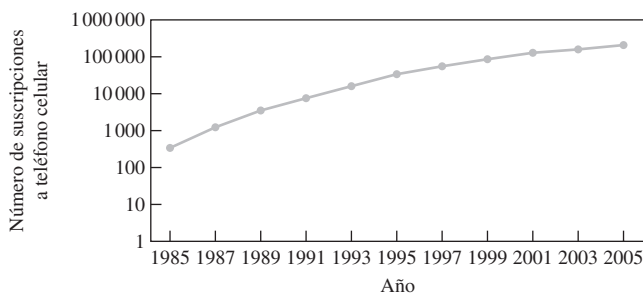


- En 2050 habrá relativamente más gente mayor y menos gente joven en la población de Estados Unidos que en 1990. Sin embargo, observe que debido al crecimiento poblacional se espera que todos los grupos de edad sean mayores en número en 2050 que en 1990.

11. a.



b.



- La gráfica en el inciso *a* nos da una mejor ilustración de la verdadera naturaleza de la tasa de cambio global. La gráfica del inciso *b* hace más sencillo ver los cambios en los primeros años, la escala exponencial es más fácil de ajustar en todos los datos.

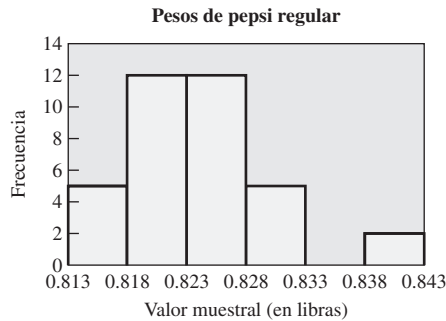
- El cambio porcentual en el IPC es más grande en 1990 y es más pequeño en 1998 o 2002, que parecen ser iguales. Los precios reales de 1991 son alrededor de 4.2% mayores que los de 1990. Los precios se han incrementado cada año, pero los aumentos casi al final del periodo son menores que los del inicio del mismo.
- El salario mínimo real en dólares no ajustados ha permanecido constante o bien se ha incrementado de manera constante desde 1955, pero el poder de compra ha

estado disminuyendo desde 1980. El poder de compra en 2006 es menor que lo que fue en 1955.

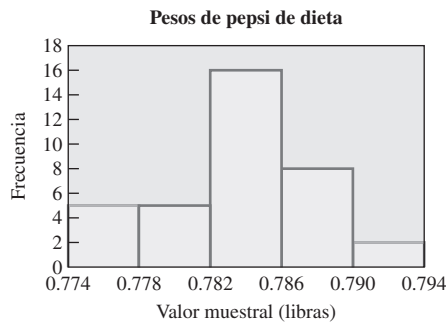
EJERCICIOS DE REPASO DEL CAPÍTULO 3

- a.** Las frecuencias son 5, 12, 12, 5, 0, 2. **b.** Las frecuencias son 5, 5, 16, 8, 2. **c.** Los pesos de pepsi regular son consistentemente mayores que los de pepsi de dieta. Los pesos de pepsi regular son mayores porque no se incluye azúcar en la pepsi de dieta.
- a.** Las frecuencias relativas son 0.139, 0.333, 0.333, 0.139, 0, 0.056. **b.** Las frecuencias acumuladas son 5, 17, 29, 34, 34, 36.

3. a.



b.

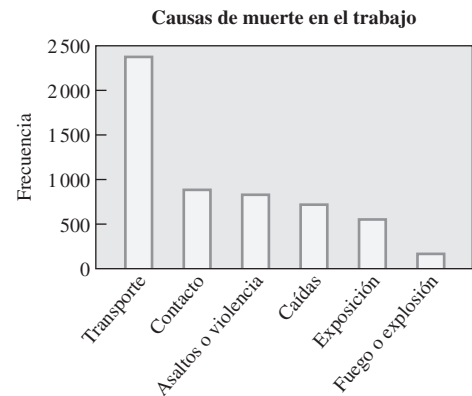


c. Las formas de los histogramas no son drásticamente diferentes, lo que indica que las distribuciones de los pesos son semejantes. Sin embargo, el rango de valores es muy diferente, indicando que los pesos de pepsi regular son considerablemente mayores que los de la pepsi de dieta.

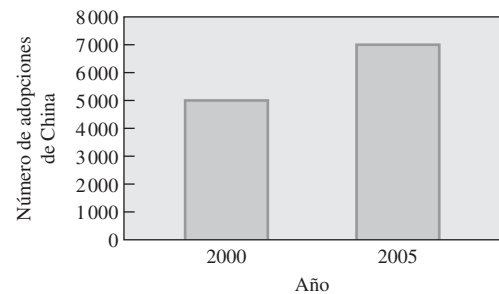
4.



- Al atraer la atención a las causas más graves de muerte, el diagrama de Pareto es más efectivo. El diagrama de Pareto también hace un mejor trabajo al mostrar la importancia relativa de las diferentes causas de muerte.



- La gráfica tiene una escala vertical que no inicia con cero, por lo que la diferencia entre las dos frecuencias está exagerada. La gráfica hace parecer que las adopciones son más del doble en 2005, pero éste no es el caso.



CUESTIONARIO DEL CAPÍTULO 3

- Gráfica de barras.
- Gráfica de múltiples barras
- Histograma
- Hay 7 valores entre 50 y 59.
- Entre todos los valores, la proporción de valores entre 50 y 59 es 0.2. Es decir, 20% de los valores están entre 50 y 59.
- 5, 10, 12, 12, 14
- Usando una escala vertical que no inicie en cero provoca que la diferencia entre las dos frecuencias se exagere.
- El valor de datos mayor es 9. La frecuencia máxima es 3 (para el valor 6).
- 25 000 en 1982; 16 652 en 2000
- Cantidades diferentes tales como precios o ingresos, la cantidad anual de gasolina no está afectada por la inflación, por lo que deben usarse las cantidades reales no ajustadas.

SECCIÓN 4.1

1. Un valor extremo en un conjunto de datos es un valor que es mucho mayor o mucho menor que casi todos los demás valores. Puesto que los valores extremos están definidos como “mucho mayores” o “mucho menores” que casi todos los demás valores, no están definidos clara y objetivamente. La determinación de valores extremos requiere de algún juicio.
3. No necesariamente. El resultado estaría distorsionado por pequeños estados con recorridos relativamente pequeños. La media debería ponderarse para tomar en cuenta las poblaciones de diferentes estados.
5. No tiene sentido. 7. Tiene sentido.
9. Los datos del ingreso con frecuencia contienen valores extremos en el lado superior, por lo que es mejor la mediana.
11. Esta distribución probablemente es muy simétrica, por lo que la media o la mediana funcionan.
13. Media: 58.3 s; mediana: 55.5 s; moda: 49 s.
15. Media: 0.188; mediana: 0.165; moda: 0.16
17. Media: 0.807 mm; mediana: 0.840 mm; moda: 0.84 mm.
19. Media: 0.9194 g; mediana: 0.9200; no existe moda.
21. **a.** Media: 157 586; mediana: 104 100. **b.** Alaska es un valor extremo superior. Sin Alaska, la media es 81 317 y la mediana es 78 650. **c.** Connecticut es un valor extremo inferior. Sin Connecticut (pero con Alaska) la media es 182 933 y la mediana es 109 050.
23. **a.** 73.75 **b.** 80 **c.** No. Una media de 80 para cinco exámenes requiere un total de 400 puntos, que no pueden alcanzarse ni con 100 en el quinto examen.
25. Con un 90, su nueva media es 81.4. La mejor calificación que podría tener es 100, lo que haría que tuviese una nueva media de 82.9. Lo peor que podría obtener es 0, lo que haría que su nueva media fuese 68.6.
27. La calificación media de sus estudiantes: 74.7; calificación mediana: 70. Por lo que depende del significado de “promedio”. Si es la media, entonces sus estudiantes están por arriba del promedio en sus calificaciones. Si es la mediana, entonces están por abajo.
29. 0.39 libras.
31. Cada estudiante está tomando tres clases con 20 inscritos en cada una y una clase con 100 inscritos, por lo que el tamaño medio del grupo para cada estudiante es $160/4 = 40$. Hay tres clases con 100 estudiantes cada una y 45 grupos con 20 estudiantes cada uno, para un total de 1 200 inscritos en 48 clases. Así, la inscripción media por clase es $1\,200/48 = 25$.
33. 0.417; es el número promedio de hit por veces al bat.

35. No, a menos que el inspector probase el mismo número de huevos en ambas granjas.
37. El resultado es no, por 600 votos contra 400 votos.
39. Media: 1.9; mediana: 2; moda: 1. La moda de 1 indica de manera correcta que los chícharos amarillo claro ocurren más que cualquier otro fenotipo, pero la media y la mediana no tienen sentido con estos datos en el nivel nominal de medida.

SECCIÓN 4.2

1. La forma de la distribución es tal que la mitad izquierda es una imagen de espejo de la mitad derecha.
3. Puesto que los estudiantes de estadística han satisfecho los requisitos y están tomando un curso del colegio, sus puntuaciones del CI probablemente tengan menos variación que las puntuaciones del CI de adultos elegidos aleatoriamente. La variación menor de puntuaciones del CI de los estudiantes de estadística da como resultado una gráfica que es más angosta y tiene menos dispersión que la gráfica de las puntuaciones del CI para adultos elegidos aleatoriamente.
5. No tiene sentido. 7. No tiene sentido.
9. Dos modas, sesgada a la izquierda, amplia variación.
11. Un solo pico, casi simétrico, variación moderada.
13. **a.** Sesgada a la derecha. **b.** 150 (la mitad de 300)
c. No; depende de la distribución.
15. **a.** Una moda de \$0 **b.** Sesgada a la derecha, ya que habrá muchos estudiantes que ganan poco o nada.
17. **a.** Una moda. **b.** Casi simétrica.
19. **a.** Dos modas. **b.** Sesgada a la derecha (suponiendo un número aproximadamente igual de patinadores en cada grupo).
21. **a.** Una moda. **b.** Sesgada a la derecha.
23. **a.** Una moda. **b.** Sesgada a la derecha.
25. **a.** Una moda. **b.** Simétrica.
27. **a.** Una moda. **b.** Sesgada a la derecha.
29. **a.** Una moda. **b.** Sesgada a la derecha.

SECCIÓN 4.3

1. La desviación estándar es una medida de cuánto los valores se desvían (o varían) de la media.
3. El enunciado es incorrecto porque define la desviación estándar como un valor que depende de los valores mínimo y máximo, mientras que la desviación estándar utiliza todos los datos.
5. No tiene sentido. 7. Tiene sentido.

9. Rango: 26.0 s; desviación estándar: 9.5 s.
11. Rango: 0.170; desviación estándar: 0.057
13. Rango: 0.280 mm; desviación estándar: 0.094 mm
15. Rango: 0.1160 g; desviación estándar: 0.0336 g
17. *Cat on a Hot Tin Roof*: rango = 10.0; desviación estándar = 2.6. *The Cat in the Hat*: rango = 3.0; desviación estándar = 0.9. Existe mucho menos variación entre las longitudes de las palabras en *The Cat in the Hat*.
19. Un día: rango: 11.0 grados; desviación estándar: 2.6 grados. Cinco días: rango = 15.0 grados; desviación estándar = 4.5 grados. Parece que hay mayor variación en los errores del pronóstico con cinco días de anticipación.
21. a. Quinto percentil. b. Percentil 69. c. Percentil 48.
23. Respuestas sólo para el conjunto 1 de datos: a. El histograma tiene una frecuencia de 7 para el valor 9; ningún otro valor. b. Valor inferior = 9; cuartil inferior = 9; mediana = 9; cuartil superior = 9; valor máximo = 9 c. 0
25. a. Facultad: media = 2, mediana = 2, rango = 4. Estudiantes: media 6.18, mediana = 6, rango = 9
b. Facultad: mínimo = 0, cuartil inferior = 1, mediana = 2, cuartil superior = 3, máximo = 4. Estudiantes: mínimo = 1, cuartil inferior = 4, mediana = 6, cuartil superior = 9, máximo = 10. c. Facultad: desviación estándar = 1.18. Estudiantes: desviación estándar = 3.03 d. Facultad: rango/4 = 1.0. Estudiantes: rango/4 = 2.25.
27. a. Primeros 7: media = 58.3, mediana = 57, rango = 4. Últimos 7: media 57.6, mediana = 56, rango = 23
b. Primeros 7: mínimo = 57, cuartil inferior = 57, mediana 57, cuartil superior = 61, máximo = 61. Últimos 7: mínimo = 46, cuartil inferior = 52, mediana = 56, cuartil superior = 64, máximo = 69 c. Primeros 7: desviación estándar = 1.9. Últimos 7: desviación estándar 7.7
d. Primeros 7: rango/4 = 1. Últimos 7: rango/4 = 5.75
29. Ordenar a la primera tienda (desviación estándar de 3 minutos).
31. Para inversiones a largo plazo, la tasa media de crecimiento mayor pagará más a largo plazo. Para inversiones a corto plazo, la menor desviación estándar es menos riesgosa.
9. Josh, Josh, Jude. 11. a. Nueva Jersey, Nebraska.
13. a. Blancos: 0.18%; no blancos: 0.54%; total 0.19%
b. Blancos: 0.16%; no blancos: 0.34%; total 0.23% c. La tasa para blancos y no blancos fue mayor en Nueva York que en Richmond, aunque la tasa global fue mayor en Richmond que en Nueva York. El porcentaje de no blancos fue significativamente menor en Nueva York que en Richmond.
15. a. Spelman tiene un mejor récord individual para juegos en casa (34.5% contra 32.1%) y para juegos fuera (75.0% contra 73.3%). b. Morehouse tiene un mejor promedio global. c. Morehouse tiene un mejor equipo, ya que, por lo general, los equipos son clasificados por récord global.
17. b. 216 personas fueron acusadas de mentir. De éstas, 18 en realidad mentían y 198 decían la verdad. c. 1 784 personas decían la verdad, de acuerdo con el polígrafo. De éstas, 1 782 en realidad decían la verdad y dos mentían; 99.9% de aquellas encontradas que decían la verdad en realidad decían la verdad.
19. Un porcentaje más alto de mujeres que de hombres fueron contratadas en ambos puestos, sugiriendo una preferencia de contratación por mujeres. Globalmente, 20% de las 200 mujeres que hicieron solicitud para posiciones de oficina (40) fueron contratadas y 85% de las 100 mujeres que solicitaron puesto de obreras (85) fueron contratadas. Así, $40 + 85 = 125$ de las $200 + 100 = 300$ mujeres que solicitaron fueron contratadas, un porcentaje de 41.7. Globalmente, 15% de los 200 hombres que solicitaron un puesto de oficina (30) fueron contratados y 75% de los 400 hombres que solicitaron un puesto de obrero (300) fueron contratados. Así, $30 + 300 = 330$ de los $200 + 400 = 600$ hombres que solicitaron un puesto fueron contratados, un porcentaje de 55.0.
21. a. En la población general, $57 + 3 = 60$ de los 20 000 en la muestra están infectados. Ésta es una tasa de incidencia de $60/20\,000 = 0.3\%$. En la población con riesgo, $475 + 25 = 500$ de los 5 000 en la muestra están infectados. Ésta es una tasa de incidencia de $500/5\,000 = 10.0\%$. b. En la categoría en riesgo, 475 de los 500 infectados tienen prueba positiva de VIH, o 95%. De aquéllos con prueba positiva, $475 + 225 = 700$ tienen VIH, un porcentaje de $475/700 = 67.9\%$. Estas dos cifras son diferentes porque miden cosas diferentes. Los 700 que dan positivo en la prueba incluyen 225 que fueron falsos positivos. Mientras que las pruebas identifican correctamente 95% de aquellos que tienen VIH, también identifica de manera incorrecta a quienes no tienen VIH. Así, sólo 67.9% de aquellos que dan prueba positiva, en realidad tienen VIH. c. En la población en riesgo, un paciente que da prueba positiva para la enfermedad tiene alrededor de 68% de posibilidad de tener la enfermedad. Esto es casi 7 veces la tasa de incidencia (10%) de la enfermedad en la categoría en riesgo. Así, la prueba es muy valiosa para

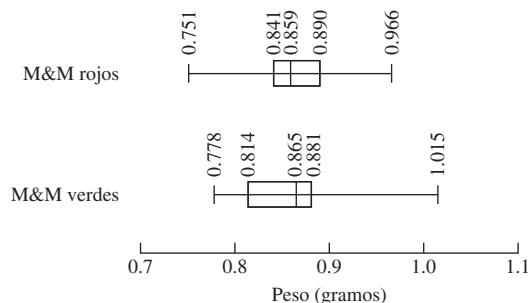
SECCIÓN 4.4

1. Un falso positivo ocurre cuando la prueba indica uso de droga para alguien que en realidad no usa drogas. Un falso negativo ocurre cuando la prueba indica que las drogas no son usadas por alguien que en realidad usa drogas.
3. La fracción de mentirosos por lo regular es pequeña, lo que significa que la fracción de personas veraces es grande. Incluso, si un porcentaje pequeño de personas veraces diesen resultados imprecisos, esto podría tomarse en cuenta para un número relativamente grande de personas.
5. Tiene sentido. 7. No tiene sentido.

identificar aquéllos con VIH. **d.** En la población general, los pacientes con prueba positiva de VIH 57 de 60, o 95% de las veces. De aquéllos con prueba positiva 57 de $57 + 997 = 1054$ en realidad tienen VIH, un porcentaje de $57/1054 = 5.4\%$. Estas dos cifras son diferentes porque miden cosas diferentes. Los 1054 que dan prueba positiva incluyen 997 que fueron falsos positivos. Mientras que la prueba identifica correctamente a 95% de quienes tienen VIH, también identifica de manera incorrecta a quienes no tienen VIH. Así, sólo 5.4% de aquellos que dan prueba positiva en realidad tienen VIH. **e.** En la población general, un paciente que da prueba positiva para la enfermedad tiene alrededor de 5.4% de posibilidades de, en realidad, tener la enfermedad. Esto es 18 veces la tasa de incidencia (0.3%) de la enfermedad en la población general. Así, la prueba es muy valiosa para identificar a aquéllos con VIH.

EJERCICIOS DE REPASO DEL CAPÍTULO 4

1. **a.** Media para rojos: 0.8635 g, media para verdes: 0.8635 g. Mediana para rojos: 0.8590 g. Mediana para verdes: 0.8650 g. **b.** Rango para rojos: 0.2150 g; rango para verdes: 0.2370 g. Desviación estándar para rojos: 0.0576 g; desviación estándar para verdes: 0.0570 g. **c.** Resumen de cinco números para rojos: 0.7510 g, 0.8250 g, 0.8590 g, 0.8975 g, 0.9660 g. Resumen de cinco números para verdes: 0.7780 g, 0.8140 g, 0.8650 g, 0.8810 g, 1.0150 g.



d. Rojos: la desviación estándar se estimó como $0.2150/4 = 0.05375$ g; la desviación estándar real es 0.0576 g. Verdes: la desviación estándar se estimó como $0.2370/4 = 0.05925$ g; la desviación estándar real es 0.0570 g. La estimación de la desviación estándar usando la regla empírica para el rango funciona razonablemente bien en ambos casos. **e.** Una comparación de los resultados de los dos conjuntos de datos sugiere que no son muy diferentes. Parece que los pesos de los M&M rojos y los M&M verdes son casi los mismos.

2. **a.** Percentil 31 **b.** 0.865 g

3. **a.** 0 **b.** Aunque ambos acumuladores tienen la misma vida media, los acumuladores con la desviación estándar menor son mejores, ya que sus vidas medias serán más cercanas a la media y pocos de ellos

dejarán varados a conductores por fallar antes de lo esperado. **c.** El valor extremo jala la media hacia arriba o hacia abajo, dependiendo si está arriba o abajo, respectivamente. **d.** El valor extremo no tiene efecto sobre la mediana. **e.** El valor extremo aumenta el rango. **f.** El valor extremo aumenta la desviación estándar.

CUESTIONARIO DEL CAPÍTULO 4

- Media.
- Desviación estándar.
- No. Es una estimación que rara vez proporciona el valor exacto.
- Cualquiera de los enunciados podría ser correcto.
- Sólo los enunciados segundo, tercero y cuarto podrían ser correctos.
- Mínimo: 30; máximo: 70
- 2
- 0
- 7.0
- Valor mínimo, primer cuartil, mediana (segundo cuartil), tercer cuartil, valor máximo.

SECCIÓN 5.1

- La palabra *normal* tiene un significado especial en estadística. Se refiere a una categoría especial de distribuciones, todas ellas con forma de campana.
- No. Los diez dígitos diferentes son igualmente probables, por lo que la gráfica de la distribución será plana, no en forma de campana.
- No tiene sentido.
- No tiene sentido.
- La distribución b no es normal. La distribución c tiene la desviación estándar mayor.
- Normal. Es común para el fabricante de un producto, tal como CD, tener una distribución que es normal. Los pesos por lo común varían arriba y abajo del peso medio casi por las mismas cantidades, por lo que la distribución tiene un pico y es simétrica.
- No es normal. Los resultados 1, 2, 3, 4, 5 y 6 todos son igualmente probables, por lo que la distribución es uniforme, no normal. Una gráfica de la distribución tenderá a ser muy plana, no en forma de campana.
- Normal. Tales mediciones físicas por lo general tienden a estar distribuidas normalmente. Un número pequeño de hombres tendrá una fuerza de sujeción muy grande y un número muy pequeño tendrá mediciones bajas. La distribución tenderá a alcanzar un máximo alrededor de la media.
- No es normal. Los tiempos de espera tenderán a estar distribuidos uniformemente.
- No es normal. Las duraciones de las películas tienen una duración mínima (cero) y no tienen duración máxima.

21. Casi normal. La desviación de la media está distribuida idénticamente alrededor de la media.

23. a. 1 b. 0.20 c. 0.80 d. 0.35 e. 0.45

25. a. 115 b. 15% c. 45% d. 15%

SECCIÓN 5.2

1. 0

3. No. La regla se aplica a distribuciones normales, pero el resultado del tiro de un dado tiene una distribución uniforme, no una distribución normal.

5. No tiene sentido. 7. Tiene sentido.

9. a. 50% b. 0.84 c. 97.5% d. 16%
e. 0.025 f. 16% g. 2.5% h. 0.84
i. 68% j. 81.5%

11. a. 68% b. 95% c. 99.7% d. 47.5%

13. 50% 15. 15.87% 17. 2.28% 19. 1.39%

21. 68.26% 23. 80.64% 25. 50.00%

27. 84.13% 29. 99.38% 31. 0.13% 33. 68.26%

35. 89.25%

37. En todos los casos, 5% de las monedas son rechazadas. Centavos: 2.44 g a 2.56 g; monedas de 5 centavos: 4.88 g a 5.12 g; monedas de 10 centavos: 2.208 g a 2.328 g; monedas de 25 centavos: 5.530 g a 5.810 g; monedas de 50 centavos: 11.060 g a 11.620 g.

39. a. 6.68% b. 48.01% c. 36.54%

41. a. 0.47% b. Aproximadamente 4% c. 30.055 a 30.745 pulg. d. La mejor estimación sería la media de las lecturas, o 30.4 pulgadas.

43. Aproximadamente 96.33%

SECCIÓN 5.3

1. No, la muestra es una muestra por conveniencia y está sujeta a un sesgo que no se esperaría con una muestra aleatoria. El estudiante no puede suponer que su muestreo por conveniencia tiene las mismas características que una muestra que fuese seleccionada aleatoriamente.

3. No. El teorema del límite central indica que la distribución de las medias muestrales será aproximadamente normal para muestras de tamaño grande, pero las muestras de tamaño 2 podrían no ser suficientes, a menos que la población original sea una distribución normal.

5. a. La media es 100 y la desviación estándar es 2.
b. La media es 100 y la desviación estándar es 1.6.
c. Con tamaños de muestra mayores (como en el inciso b), las medias tienden a estar más juntas, de modo que tienen menos variación, lo que da como resultado una desviación estándar menor.

7. a. La media es 6.5 y la desviación estándar es 0.384.

b. La media es 6.5 y la desviación estándar es 0.345.

c. Con tamaños de muestra mayores (como en el inciso b), las medias tienden a estar más juntas, de modo que tienen menos variación, lo que da como resultado una desviación estándar menor.

9. 40%, 4% 11. 29.6%, 99.96%

13. a. 0.35% b. No, parece que la media es mayor que 12.00 onzas, pero los consumidores no están siendo engañados ya que las latas están sobrellenándose no subllenándose.

15. a. 57.93% de posibilidad. b. 97.72% de posibilidad. c. Aunque el ancho medio de la cabeza de 100 hombres es muy probable que sea menor de 6.2 pulgadas, podría haber muchos individuos hombres que no podrían usar los cascos porque tuviesen anchos de cabeza mayor que 6.2 pulgadas. Con base en los resultados del inciso a, estos cascos no se ajustan para alrededor de 42% de los hombres.

17. a. 52.91% b. Aproximadamente una posibilidad de 73%. c. El inciso a, porque los asientos serán ocupados por mujeres de manera individual, no grupos de mujeres.

19. a. 16 (5.74% de 280). b. Las puntuaciones estándar z están más allá de las incluidas en la tabla 5.1, pero la tabla 5.1 muestra que es muy poco probable (menos de 0.04% de posibilidad) que la media caiga entre los límites de 5.550 g y 5.790 g. c. El inciso a, ya que las monedas rechazadas podrían dar como resultado ventas perdidas y menos utilidades.

21. a. La desviación estándar de la distribución de medias muestrales es $\sigma/\sqrt{n} = \sigma/10$. b. La desviación estándar de la distribución de medias muestrales es aproximadamente $\sigma/\sqrt{1000} \approx \sigma/32$, que es alrededor de un tercio de la desviación estándar del inciso a. c. La desviación estándar de la distribución de medias muestrales disminuye conforme n aumenta.

EJERCICIOS DE REPASO DEL CAPÍTULO 5

1. a. Puesto que los 38 resultados son igualmente probables, la distribución es uniforme, no normal. b. Los pesos de una población homogénea, tal como los perros Golden Retriever, por lo común tienen una distribución normal. c. Aproximadamente normal. Las medidas físicas y psicológicas humanas tienden a estar distribuidas normalmente. (Sin embargo, la distribución podría ser ligeramente sesgada, ya que no puede tener tiempos de reacción negativos, pero hay poca gente con muy altos tiempos de reacción).

2. a. 95% b. 99.7% c. Sí, puesto que tal calificación alta sólo ocurre alrededor de 0.3% del tiempo.

3. **a.** Percentil 90 **b.** 1.29 **c.** No, el dato está a menos de 2 desviaciones estándar de la media. **d.** 0.0065
e. La temperatura es inusual; está a menos de 2 desviaciones estándar de la media. **f.** 99.22 **g.** 97.18
h. Poco menos de 0.01% **i.** La media muestral está a 6.6 desviaciones estándar de la media; la posibilidad de seleccionar esa muestra es extremadamente pequeña. La media supuesta (98.60) podría ser incorrecta.

CUESTIONARIO DEL CAPÍTULO 5

1. **b** y **e** son correctas. 2. 95% 3. 50 4. 0.5
 5. 2 6. -1 7. 50% 8. 2.28% 9. 2.28%
10. Inciso c.

SECCIÓN 6.1

1. La afirmación no es válida, porque es fácil obtener 11 mujeres en 20 nacimientos por azar.
3. No. La significancia estadística al nivel 0.05 significa que existe menos de 0.05 de probabilidad de que el resultado ocurra por azar, pero una probabilidad menor a 0.05 no necesariamente significa que exista menos de 0.01 de posibilidades.
5. No tiene sentido. 7. Tiene sentido.
9. No es estadísticamente significativa. Los resultados son cercanos a los 250 cruces esperados, por lo que podrían haber sido resultado del azar.
11. Es estadísticamente significativo. Con seis resultados posibles, esperamos que alrededor de 10 de los 60 resultados sean 3, así que ningún 3 es muy poco probable que ocurra por azar.
13. Es estadísticamente significativa. Es muy poco probable que cuando 20 adultos se seleccionan aleatoriamente, todos sean mujeres.
15. No es estadísticamente significativa. El resultado podría haber ocurrido por azar.
17. Aunque el tamaño de la muestra es pequeño, 21% de mejora en el millaje es significativo.
19. Con 325 bebés, el número de niñas usualmente sería alrededor de 162, por lo que el resultado de 295 niñas se sale sustancialmente del resultado que se esperaba por el azar. Los resultados parecen ser estadísticamente significativos.
21. **a.** Si fuesen seleccionadas 100 muestras, la temperatura media sería 98.20 o menos en 5 o menos de las muestras.
b. La selección de una muestra con una media pequeña es extremadamente poco probable y no se esperaría por el azar.
23. Este resultado no es significativo al nivel 0.05 ya que la probabilidad de ocurrir por el azar cuando no hay una mejora real es mayor que 0.05.

SECCIÓN 6.2

1. $P(A)$ representa la probabilidad de que ocurra el evento A . $P(\text{no } A)$, o $P(\bar{A})$, es la probabilidad de que el evento A no ocurra.
3. El razonamiento es incorrecto ya que supone que los dos resultados de llover o no llover son igualmente probables, pero no son igualmente probables.
5. Tiene sentido. 7. Tiene sentido.
9. No tiene sentido.
11. **a.** Resultado. **b.** Evento, ya que puede ocurrir de tres formas diferentes. **c.** Resultado. **d.** Evento, ya que puede ocurrir de tres formas diferentes.
e. Evento, ya que puede ocurrir de seis formas diferentes.
f. Evento, ya que puede ocurrir de seis formas diferentes.
13. $1/3$, suponiendo que el dado no está cargado y los resultados son igualmente probables.
15. $2/38$, o $1/19$, suponiendo que los resultados son igualmente probables.
17. $1/365$, suponiendo que los nacimientos en los 365 días son igualmente probables.
19. $1/2$, o 0.5 21. $4/6$, o $2/3$. 23. $36/38$, o $18/19$
25. 0.45 27. 0.720 29. 0.22, 0.33, 0.44, 0.56
31. **a.** $1/8 = 0.125$ (MMM) **b.** $3/8 = 0.375$ (HHM, HHM, HHH) **c.** $1/8 = 0.125$ (MHH) **d.** $7/8 = 0.875$ (MMM, HMM, MHM, MMH, HHM, HMM, MHH)
e. $4/8 = 0.5$ (HHM, HHM, MHH, HHH)
33. 0.40 35. $P(\text{éxito}) = 0.86$
37. La probabilidad de que conozca a una persona al azar de más de 65 años será de $34.7 \text{ millones} / 281 \text{ millones} = 0.123$ en el año 2000 y será $78.9/394 = 0.200$ en 2050. Así, sus posibilidades serán mayores en 2050.
39. **a.** Los resultados al lanzar cuatro monedas legales:

Moneda 1	Moneda 2	Moneda 3	Moneda 4	Resultado	Probabilidad
H	H	H	H	HHHH	1/16
H	H	H	T	HHHT	1/16
H	H	T	H	HHTH	1/16
H	H	T	T	HHTT	1/16
H	T	H	H	HTHH	1/16
H	T	H	T	HTHT	1/16
H	T	T	H	HTTH	1/16
H	T	T	T	HTTT	1/16
T	H	H	H	THHH	1/16
T	H	H	T	THHT	1/16
T	H	T	H	THTH	1/16
T	H	T	T	THTT	1/16
T	T	H	H	TTHH	1/16
T	T	H	T	TTHT	1/16
T	T	T	H	TTTH	1/16
T	T	T	T	TTTT	1/16

b. La distribución de probabilidad para el número de caras en el lanzamiento de cuatro monedas legales:

Resultado	Probabilidad
4 caras (0 cruces)	1/16
3 caras (1 cruz)	4/16
2 caras (2 cruces)	6/16
1 cara (3 cruces)	4/16
0 caras (4 cruces)	1/16
Total	1

c. $6/16 = 0.375$ d. $15/16 = 0.9375$ e. 2 caras (la probabilidad es 0.375)

SECCIÓN 6.3

- Si algún proceso se repite en muchos ensayos, la proporción de ensayos en la que algún evento ocurrirá será cercana a la probabilidad del evento. Cuando el número de ensayos aumenta, la proporción de ensayos en los que el evento ocurre se hace más cercana a la probabilidad del evento.
- Su estrategia de apuesta es imprudente ya que el casino continúa teniendo una ventaja. Su desempeño pasado no afecta a eventos futuros. Su razonamiento no es correcto. Aunque su proporción de apuestas ganadas es probable que aumente, aunque él continúe perdiendo más dinero. Esta falla en el razonamiento ha causado que muchos jugadores pierdan grandes cantidades de dinero.
- Tiene sentido. 7. Tiene sentido.
- No. Usted no debe esperar obtener exactamente 5 000 caras, ya que la probabilidad de ese resultado particular es extremadamente pequeña. La proporción de caras debe aproximarse a 0.5 conforme el número de lanzamientos aumenta.
- El valor esperado para usted en un juego es $(\$5 \times 0.25) + (-\$1 \times 0.75) = \$1.25 - 0.75 = \0.50 . Puesto que debe ganar \$5 o perder \$1 en el primer juego, no puede ganar el valor esperado de \$0.50. Si juega 100 partidas, debe esperar ganar $100 \times \$0.50 = \50 .
- 0.94, 0.74
- 15 minutos 17. $-\$0.78$, $-\$285$ 19. -7.07 centavos, -1.4 centavos
- a. **Decisión 1** Opción A: valor esperado = \$1 000 000; Opción B: valor esperado = \$1 140 000. **Decisión 2** Opción A: valor esperado = \$110 000; Opción B: valor esperado = \$250 000 b. Las respuestas no son consistentes con los valores esperados en la decisión 1, pero lo son con la decisión 2. c. Parece que la gente elige el resultado cierto (\$1 000 000) en la decisión 1.

- a. 0.5, 0.5 b. Pérdida de \$10 c. Pérdida de \$16 d. Pérdida de \$20 e. 45%, 46%, 48%. El porcentaje de números pares se aproxima a 50%, pero la diferencia aumenta entre los números pares y los números impares. f. 60 números pares.

SECCIÓN 6.4

- Puesto que son expresados como los números de nacimientos por cada 1 000 personas en la población, las tasas de natalidad pueden compararse directamente. Los números reales de nacimientos deben considerarse en el contexto de los tamaños de las poblaciones de los países, así que una comparación requeriría los números de nacimientos junto con los tamaños de la población, e incluso entonces la comparación no sería fácil a consecuencia de los números involucrados.
- La esperanza de vida es el número de años que una persona con una edad dada puede, en promedio, esperar vivir. Una persona de 30 años tendrá una esperanza de vida menor que una persona de 20 años. La persona de 30 años no esperaría vivir más años adicionales que una persona de 20 años.
- No tiene sentido. 7. Tiene sentido.
- 1995: 0.0208; 2000: 0.0102; 2004: 0.0013. El año 2004 fue el más seguro ya que tuvo el número más bajo de accidentes mortales por cada 1 000 salidas.
- 1995: 3.1; 2000: 1.4; 2004: 0.2. El año 2004 fue el más seguro ya que tuvo el número más bajo de accidentes mortales por cada 10 millones de pasajeros.
- 58.8 años 15. 8.3 muertes por cada 10 000 personas.
- a. Utah: 137; Maine: 38 b. Utah: 20.8; Maine: 10.7
- a. 4 230 000 b. 2 520 000 c. 1 710 000 d. 0.0057; 0.57%

SECCIÓN 6.5

- La ocurrencia de uno de los eventos no afecta la probabilidad de otro evento.
- Muestreo con reemplazo. El segundo resultado es independiente del primero.
- No tiene sentido. 7. No tiene sentido.
- $1/8$ o 0.125
- $1/260$; la contraseña es demasiado corta para ser efectiva.
- a. 0.0039 b. 0.00098 c. 0.125 d. 0.0625 e. Hay 60 canciones igualmente probables disponibles para cada selección. No importa cuál canción se toque primero, la probabilidad de que la siguiente sea la misma es $1/60 = 0.0167$.

15. P (declararse culpable o enviar a prisión) = $1\,014/1\,028 = 0.986$

17. 0.865 19. 0.381 21. 0.410 23. 0.920

25. 0.0195

27. a. 0.733 b. 1 c. 0.643 d. 0.22

29. a. 0.5625 b. 0.375 c. 0.0625

d.

Evento	Probabilidad
AA	0.5625
Aa	0.1875
aA	0.1875
aa	0.0625

e. 0.9375

31. a. 0.518 b. 0.491

EJERCICIOS DE REPASO DEL CAPÍTULO 6

1. 85/99, o 0.859 2. 83/99, o 0.838

3. 88/99, o 0.889 4. 96/99, o 0.970

5. 0.702 6. 0.736

7. a. 0.73 b. 0.073 c. 1.35 d. 0.0014; debe dudar del rendimiento establecido.

8. a. 0.10 b. No, la probabilidad es mayor que 0.05.

9. a. $1/38 = 0.026$ b. Sí, la probabilidad es menor que 0.05

10. a. 0.14 b. No, la probabilidad es mayor que 0.05.

CUESTIONARIO DEL CAPÍTULO 6

1. $1/6$ 2. $1/216$, o 0.00463 3. 0.55 4. 7.5

5. Sí 6. 0.502 7. 0.141 8. 0.613

9. 0.251 10. 0.123

SECCIÓN 7.1

1. Una correlación existe entre dos variables cuando valores grandes de una tienden a estar consistentemente asociados con valores grandes del otro, o cuando valores grandes de uno de manera consistente tienden a estar asociados con valores pequeños de otra. El significado del término *correlación* en estadística es mucho más específico que su significado en uso general.

3. Los puntos son muy cercanos a un patrón lineal (o línea recta), y tienen alturas mayores entre más a la derecha se encuentren.

5. No tiene sentido. 7. No tiene sentido.

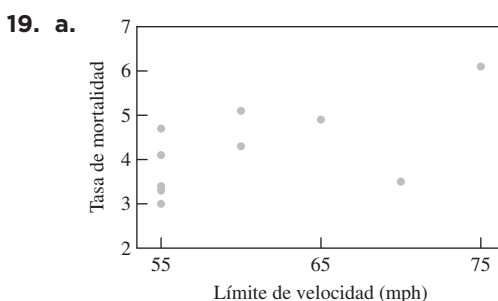
9. Correlación positiva porque mujeres más altas tienden a pesar más.

11. Correlación negativa porque automóviles que pesan más tienden a usar más gasolina, por lo que el consumo de gasolina en millas por galón será menor.

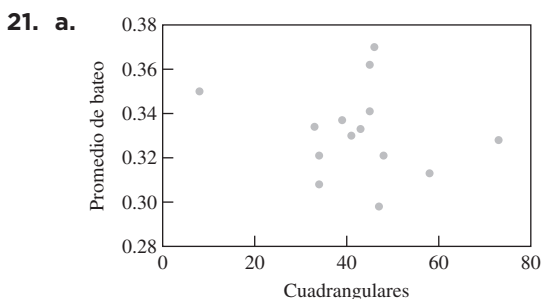
13. Correlación positiva porque correr el maratón más rápido (en menos tiempo) dará como resultado un orden menor de terminación.

15. Las variables no están correlacionadas.

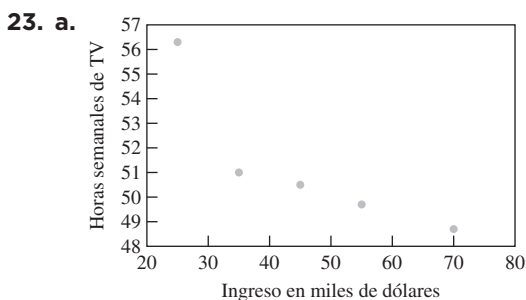
17. Hay una fuerte correlación positiva, con el coeficiente de correlación aproximadamente 0.8. Gran parte de esta correlación se debe al hecho de que una fracción grande del grano producido es usado para ganado.



b. Existe una moderada correlación positiva ($r = 0.59$, exactamente). c. Con la excepción de Gran Bretaña, los límites de velocidad por lo general están asociados con altas tasas de mortalidad. Éstas también están influidas por otros factores además de los límites de velocidad.

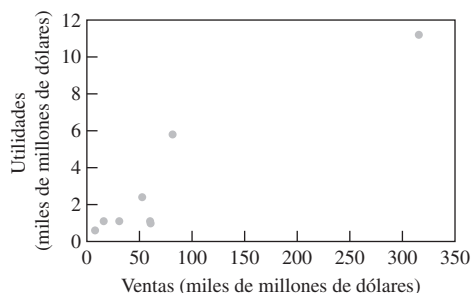


b. Existe una moderada correlación negativa ($r = -0.29$ exactamente). c. No



- b.** Existe una fuerte correlación negativa entre el ingreso y el número de horas semanales de TV ($r = -0.86$ exactamente). **c.** Las familias con más ingresos tienen oportunidades de hacer otras cosas. No.

25.a.



- b.** Hay una fuerte correlación entre las ventas e ingresos ($r = 0.92$ exactamente). En este caso la fuerte correlación es muy afectada por el valor extremo. **c.** Ventas más altas no necesariamente se traducen en ingresos más altos. Algunas compañías tienen grandes gastos, llevando a una disminución de ingresos.

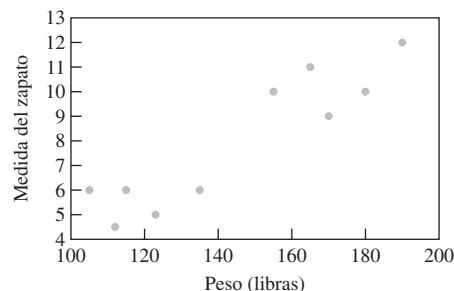
27. Las variables x y y aparecen simétricamente en la fórmula (intercambiar x y y no cambia la fórmula).

SECCIÓN 7.2

1. Si existe una correlación significativa entre dos variables existe una asociación estadística, pero esta asociación no implica que una de las variables de alguna variable sea la causa o efecto directo de la otra variable.
3. Los valores extremos están muy alejados de casi todos los otros valores en un conjunto de datos. Los valores extremos podrían hacernos parecer que existe una correlación significativa cuando no existe, o podría ocultar una correlación real.
5. No tiene sentido. 7. No tiene sentido.
9. Existe una correlación positiva debida probablemente a una causa subyacente común. Muchos crímenes son cometidos por pistolas que no están registradas.
11. Existe una correlación positiva que se debe a una causa directa. Los peajes a lo largo de la autopista de Massachusetts están basados en la distancia recorrida.
13. Existe una correlación positiva que probablemente se deba a una causa común, tal como el aumento general en el número de automóviles y el tráfico.
15. Existe una correlación negativa debida probablemente a una causa directa. Conforme los precios de la gasolina aumentan en grandes cantidades, las personas no pueden permitirse manejar mucho, por lo que reducen costos manejando menos.
17. **a.** El valor extremo es el punto de la parte superior izquierda (0.4, 1.0). Sin el valor extremo, el coeficiente

de correlación es 0.0. **b.** Con el valor extremo, el coeficiente de correlación es -0.58

19. **a.** El coeficiente de correlación real es $r = 0.92$, que es significativo al nivel 0.01, por lo que hay una fuerte correlación entre el peso y la medida del zapato.



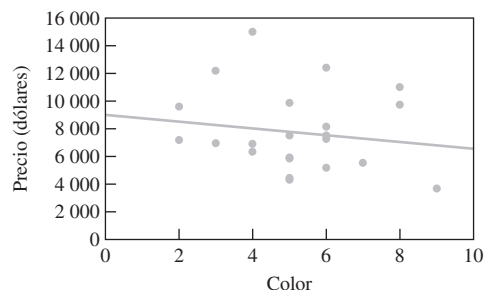
- b.** Dentro de cada uno de los dos grupos, la correlación es menos fuerte que la correlación global.

21. **a.** El coeficiente de correlación real es $r = 0.77$, que es significativo al nivel 0.01, indicando una fuerte correlación. **b.** Los 16 puntos de la derecha corresponden a países relativamente pobres, como Uganda. Los puntos restantes corresponden a países relativamente ricos, como Suecia. **c.** Parece haber una correlación negativa entre las variables para los países pobres y una correlación positiva para los países ricos.

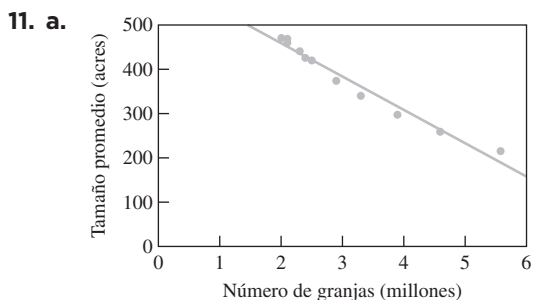
SECCIÓN 7.3

1. Una recta de mejor ajuste (o recta de regresión) es una recta en un diagrama de dispersión que está más cerca de los puntos que cualquier otra posible recta (de acuerdo con la medida estadística estándar de cercanía). Una recta de mejor ajuste es útil para pronosticar el valor de una variable dando algún valor de la otra variable.
3. r^2 es el cuadrado del coeficiente de correlación y es la proporción de la variación en una variable que puede explicarse por la recta de mejor ajuste.
5. No tiene sentido. 7. No tiene sentido.

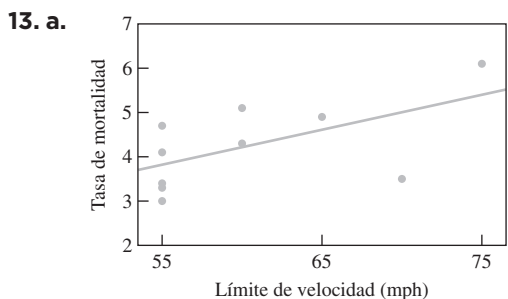
9.a.



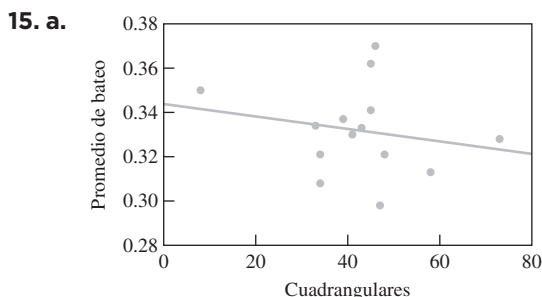
- b.** Real $r = -0.16$; $r^2 = 0.026$. Alrededor de 3% de la variación en el precio puede explicarse mediante la recta de mejor ajuste. **c.** La recta de mejor ajuste no debe usarse para hacer pronósticos.



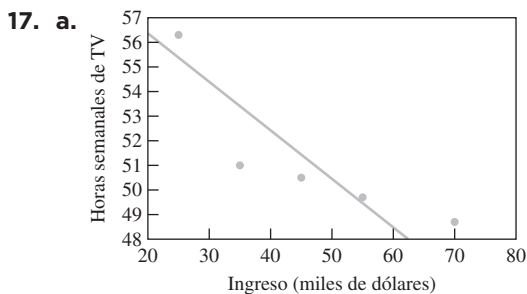
b. Real $r = -0.99$; $r^2 = 0.97$. Alrededor de 97% de la variación en el tamaño de las granjas puede explicarse mediante la recta de mejor ajuste. **c.** La recta de mejor ajuste podría usarse para hacer predicciones dentro del rango del número de granjas incluido en los datos. Puesto que parece haber una ligera curvatura en los puntos, las predicciones no deben hacerse fuera de ese rango.



b. Real $r = 0.59$; $r^2 = 0.35$; 35% de la variación puede ser tomada en cuenta por la recta de mejor ajuste. **c.** (75, 6.1) y (70, 3.5) ambos son posibles valores extremos, el primero porque está lejos de la mayoría de los datos y el último porque la tasa de mortalidad es menor de lo que podría esperarse considerando el resto de los datos. Puesto que un punto está por arriba de la recta de mejor ajuste y otro está por abajo, el efecto neto de los dos puntos probablemente es que se cancelen. **d.** El valor de r es demasiado pequeño para predicciones basadas en la recta de mejor ajuste, para considerarse confiables.

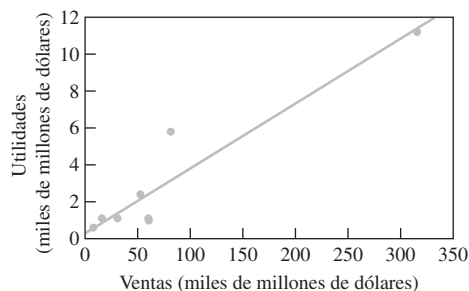


b. Real $r = -0.29$; $r^2 = 0.08$. Sólo 8% de la variación puede ser tomada en cuenta por la recta de mejor ajuste. **c.** (.366, 49) es un valor extremo, ya que está muy lejos de los otros puntos. **d.** Las predicciones basadas en la recta de mejor ajuste no son confiables.



b. Real $r = -0.86$; $r^2 = 0.74$; 74% de la variación puede tomarse en cuenta por la recta de mejor ajuste. **c.** (25 000, 56.3) es un valor extremo. La recta de mejor ajuste tendría una menor pendiente descendente si ese punto fuera eliminado. **d.** Las predicciones basadas en la recta de mejor ajuste podrían ser confiables, pero la presencia del valor extremo y su efecto sobre la recta de mejor ajuste hace cuestionable la confiabilidad.

19. a. Real $r = 0.92$; $r^2 = 0.84$; 84% de la variación puede ser tomada en cuenta por la recta de mejor ajuste.



b. (81.5, 5.8) es un valor extremo. **c.** Las predicciones basadas en la recta de mejor ajuste podrían ser confiables, pero el fuerte efecto del valor extremo en la determinación de la recta de mejor ajuste hace cuestionable la confiabilidad. Se necesitan más datos.

SECCIÓN 7.4

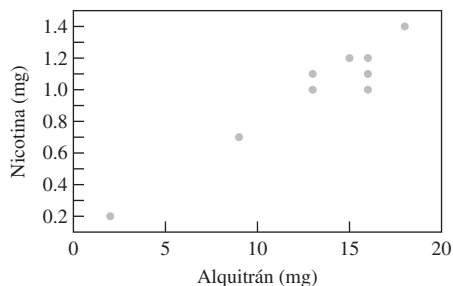
- (1) La correlación es el resultado de una coincidencia; (2) la correlación es debida a una causa común; (3) la correlación es debida a una influencia directa de una de las variables sobre la otra.
- Una variable de confusión es una variable que no está incluida en el análisis pero que afecta a la variable incluida en el análisis. No incluir o tomar en cuenta una variable de confusión podría causar que un investigador pierda una causalidad subyacente.
- No tiene sentido. **7.** No tiene sentido.
- La relación causal es válida. Cuando la velocidad del palo aumenta, más fuerza es aplicada a la pelota de golf, de modo que viaja más rápido.
- La relación causal es válida. El alcohol es un depresor del sistema nervioso central y tiene varios efectos que incluyen una disminución en la velocidad de reacción. Ésta

es una razón importante por la que manejar y beber es muy peligroso.

13. Directriz 1, directrices 2 y 5; directrices 3 y 5. Los dolores de cabeza están asociados con días de trabajo de alguna manera. Los dolores de cabeza no están asociados con coca o la cafeína de la coca. Los dolores de cabeza quizá son el resultado de mala ventilación en el edificio.
15. Fumar sólo puede aumentar el riesgo ya presente.
17. Éste fue un estudio de observación. Un hijo tardío refleja una causa subyacente. Aunque es posible que las conclusiones sean correctas, hay otras explicaciones posibles para los hallazgos. Por ejemplo, es posible que las mujeres más jóvenes vivieron durante una época en que era menos probable tener bebés después de los 40 (por elección propia). Es posible para ellas vivir hasta los 100.
19. La disponibilidad en sí misma no es una causa. Las condiciones sociales, económicas o personales hacen que los individuos usen las armas de fuego.

EJERCICIOS DE REPASO DEL CAPÍTULO 7

1. La correlación es fuerte y positiva; 92.5%.
2. La correlación es fuerte y positiva; 82.6%
3. La correlación es fuerte y positiva; 95.8%
4. Puesto que el patrón de puntos es cercano al patrón de una línea recta, el diagrama de dispersión indica que hay una correlación entre las variables.



5. Los datos consistirán de la tasa de mortalidad en un vecindario y la distancia entre el vecindario y los cables de electricidad para muchos vecindarios. Podría ser posible establecer una correlación entre los cables de electricidad y las muertes por leucemia, pero sería muy difícil establecer una relación causal.
6. Los puntos en el diagrama de dispersión están en una línea recta con pendiente negativa (desciende hacia la derecha).
7. La correlación sola nunca implica causalidad y, en este caso, seguramente más visitas al dentista no causan mayores ingresos. Los hogares con mayores ingresos disponibles pueden permitirse más visitas al dentista o

pueden permitirse un seguro dental que cubra el costo de las visitas.

8. Las variables que afectan el valor de una casa podrían incluir su ubicación, tamaño, antigüedad, condición y tamaño del terreno. Con frecuencia la ubicación es citada como el factor más importante, con consideraciones que incluyen cercanía a escuelas, comercios e iglesias. La edad del propietario anterior no estaría relacionada con el valor.
9. Los datos que fueron recolectados no estaban correlacionados. Pero es posible que las variables representadas por los datos estén relacionadas de alguna manera no lineal, por ejemplo, el diagrama de dispersión forma una curva en lugar de una línea recta.
10. El coeficiente de correlación es -0.056 , pero cualquier valor positivo o negativo cercano a cero es razonable.

CUESTIONARIO DEL CAPÍTULO 7

1. -1.1 2. c, d y e 3. c 4. El valor real de r es -0.934 , pero cualquier valor entre -0.5 y -0.99 es una respuesta razonable.
5. Sí 6. Falso 7. Falso 8. Verdadero
9. Falso 10. Verdadero

SECCIÓN 8.1

1. Uniforme; aproximadamente normal.
3. $\bar{x}$ denota la media de una muestra, mientras que μ denota la media de una población.
5. No tiene sentido. 7. Tiene sentido.
9. La mejor estimación de la proporción poblacional para el noreste es la proporción muestral: 0.19. La proporción muestral dada no es probable que sea una buena estimación de la proporción de la población de niños en Estados Unidos. Existen factores regionales que podrían tener una fuerte influencia en la proporción en partes diferentes del país.
11. a. 9.5 desviaciones estándar por arriba de la media ($z = 9.5$)
b. La probabilidad es muy pequeña, tal como 0.01%
c. No. Parece que las latas están siendo llenadas con una cantidad que es mayor que 12.00 onzas.
13. a. 0.528 b. 0.49 c. La proporción muestral es ligeramente muy baja.
15. 783 personas. Es probable que la muestra más grande proporcione una estimación más confiable para la estimación de la proporción poblacional.
17. a. -0.67 desviación estándar por abajo de la media ($z = -0.67$) b. 0.253
19. a. 1.0, 1.5, 3.0, 1.5, 2.0, 3.5, 3.0, 3.5, 5.0 b. 2.7
c. Sí; sí.

21. Muestras de tamaño $n = 1$

Muestra	Media
A	52
C	5
G	60
M	33
O	97
Media	49.4

Muestras de tamaño $n = 2$

Muestra	Media
AC	28.5
AG	56.0
AM	42.5
AO	74.5
CG	32.5
CM	19.0
CO	51.0
GM	46.5
GO	78.5
MO	65.0
Media	49.4

Muestras de tamaño $n = 3$

Muestra	Media
ACG	39.00
ACM	30.00
ACO	51.33
AGM	48.33
AGO	69.67
AMO	60.67
CGM	32.67
CGO	54.00
CMO	45.00
GMO	63.33
Media	49.4

Muestras de tamaño $n = 4$

Muestra	Media
ACGM	37.50
ACGO	53.50
ACMO	46.75
AGMO	60.50
CGMO	48.75
Media	49.4

Muestras de tamaño $n = 5$

Muestra	Media
ACGMO	49.4
Media	49.4

SECCIÓN 8.2

1. Tenemos 95% de confianza en que los límites de 183.3 g y 298.5 g en realidad contienen la media poblacional verdadera del nivel de colesterol de todas las mujeres. Esperamos que 95% de tales muestras den como resultado límites de intervalos de confianza que contengan la media poblacional.
3. Con frecuencia, los medios omiten la referencia del nivel de confianza, que comúnmente es 95%. La palabra “media” debe usarse en lugar de la palabra “promedio”.
5. No tiene sentido. 7. No tiene sentido.
9. Margen de error: 2.0. $73.0 < \mu < 77.0$
11. Margen de error: \$1 469. $\$44\,736 < \mu < \$47\,674$
13. 49 15. 484 17. 185 19. 225
21. 5.639 g; $5.619\text{ g} < \mu < 5.659\text{ g}$
23. 5.15 años; $5.10\text{ años} < \mu < 5.20\text{ años}$
25. $146\text{ lb} < \mu < 221\text{ lb}$
27. a. 3.1 b. 1.4 c. 3.1 d. $2.6 < \mu < 3.6$
e. La muestra es pequeña y estaba concentrada en una vecindad, por lo que es poco probable que sea representativa de todas las familias estadounidenses.

SECCIÓN 8.3

1. Tenemos 95% de confianza que los límites de 0.457 y 0.551 en realidad contienen la proporción poblacional verdadera. Esperamos que 95% de tales muestras den como resultado límites de intervalos de confianza que contengan la proporción poblacional.
3. Con frecuencia los medios omiten referencia al nivel de confianza, que comúnmente es 95%.
5. No tiene sentido. 7. No tiene sentido.
9. Margen de error: 0.0866. $0.163 < p < 0.337$
11. Margen de error: 0.0257. $0.202 < p < 0.254$
13. 10000 15. 278 17. $0.140 < p < 0.160$
19. $E = 0.021$; $0.779 < p < 0.821$
21. a. $0.153 < p < 0.191$ b. $0.118 < p < 0.146$
c. $0.215 < p < 0.243$
23. a. 0.98 b. $0.966 < p < 0.994$ c. Para la muestra que sea una muestra aleatoria, todas las películas deben tener una oportunidad igual de ser elegidas.
25. El intervalo de confianza del 95% es $0.538 < p < 0.587$. Es razonable una predicción de victoria.
27. Sondeo 1: $0.494 < p < 0.546$; Sondeo 2: $0.494 < p < 0.534$
Sondeo 3: $0.498 < p < 0.532$. Puesto que todos estos

intervalos de confianza incluyen valores menores que 0.5. Martínez no puede estar confiado de ganar una mayoría.

29. a. $0.50 < p < 0.58$ b. 621

EJERCICIOS DE REPASO DEL CAPÍTULO 8

1. a. $0.008 < p < 0.056$ b. Hay una confianza de 95% que los límites de 0.008 y 0.056 contienen los valores verdaderos de la proporción poblacional. Si tales muestras de tamaño 221 fuesen seleccionadas aleatoriamente muchas veces, los intervalos de confianza resultantes contendrían la proporción poblacional verdadera en 95% de esas muestras. c. 0.0237 d. La distribución será aproximadamente normal o en forma de campana. e. Estarían más cercanos. f. La proporción poblacional es un valor fijo, no una variable aleatoria, y está contenida dentro de los límites del intervalo de confianza o no está allí, por lo que no hay una probabilidad involucrada con él.
2. a. 114 b. Si el tamaño de la muestra es mayor que lo necesario, el intervalo de confianza será mejor en el sentido que será más angosto de lo necesario. Si el tamaño de la muestra es más pequeño de lo necesario, el intervalo de confianza será peor en el sentido que será más amplio de lo que debe ser.
c. Puesto que los estudiantes universitarios son un grupo más homogéneo que la población como un todo, la desviación estándar para estudiantes universitarios es probable que sea menor a 16. Si usamos el valor real que es menor a 16, el tamaño de la muestra será menor que 114.
3. a. $6.4 < \mu < 7.8$ b. Hay 95% de confianza de que los límites de 6.4 y 7.8 contienen el valor verdadero de la media poblacional. Si tales muestras de tamaño 50 fuesen seleccionadas aleatoriamente contendrían la media poblacional verdadera en 95% de esas muestras. c. 0.707 d. 400
4. a. 2500 b. No. Su tamaño de muestra será suficientemente grande, pero tiene un alto potencial para estar sesgada. La gente que rehúsa responder podría constituir un segmento de la población con una perspectiva diferente y esa perspectiva será excluida de manera incorrecta de la muestra.

CUESTIONARIO DEL CAPÍTULO 8

1. Aproximadamente normal o forma de campana.
2. Aproximadamente normal o forma de campana.
3. La proporción muestral 4. $0.52 < p < 0.68$ 5. b
6. 0.02 7. 20 8. $4.45 < \mu < 4.75$
9. $0.400 < p < 0.800$
10. Los límites del intervalo de confianza deben ser proporciones entre 0 y 1, pero los datos no.

SECCIÓN 9.1

1. Una prueba de hipótesis es un procedimiento estándar para probar una afirmación acerca del valor de un parámetro poblacional.
3. $\mu \neq 98.6, \mu < 98.6, \mu > 98.6$
5. No tiene sentido. 7. No tiene sentido.
9. No tiene sentido. 11. No tiene sentido.
13. a. El aumento es relativamente pequeño, de modo que no parece ser estadísticamente significativo. b. El aumento es muy grande, así que parece ser estadísticamente significativo. c. Las respuestas variarán, pero un aumento de alrededor de tres o más galones parecería ser estadísticamente significativo.
15. $H_0: \mu = 1 \text{ L}, H_a: \mu < 1 \text{ L}$. No hay evidencia muestral suficiente para respaldar la afirmación de que la cantidad media de cola es menor que 1 L. Hay evidencia suficiente para respaldar la afirmación de que la cantidad media de cola es menor que 1 L.
17. $H_0: p = 0.7, H_a: p > 0.7$. No hay evidencia muestral suficiente para respaldar la afirmación de que la proporción de niñas es mayor que 0.7. Hay evidencia suficiente para respaldar la afirmación de que la proporción de niñas es mayor que 0.7.
19. $H_0: \mu = 10 \text{ onzas}, H_a: \mu \neq 10 \text{ onzas}$. No hay evidencia muestral suficiente para justificar el rechazo de la afirmación de que la cantidad media de café es igual a 10 onzas. Hay evidencia muestral suficiente para justificar el rechazo de la afirmación de que la cantidad media de café es igual a 10 onzas.
21. $H_0: p = 0.5, H_a: p > 0.5$. No hay evidencia muestral suficiente para respaldar la afirmación de que la proporción de estudiantes es mayor que 0.5. Hay evidencia suficiente para respaldar la afirmación de que la proporción de estudiantes es mayor que 0.5.
23. No. El valor P de 0.382 es mayor que el nivel de significancia de 0.05, y este valor alto sugiere que hay altas posibilidades de obtener 48 o menos hombres por azar.
25. No. El valor P de 0.028 es mayor que el nivel de significancia de 0.01, y este valor alto sugiere que hay altas posibilidades de obtener 40 o menos hombres por azar.
27. Sí. El valor p de 0.002 es menor que el nivel de significancia de 0.01 y este valor bajo sugiere que hay pocas posibilidades de obtener 35 o menos hombres por azar.

SECCIÓN 9.2

1. n representa el tamaño de la muestra, $\bar{x}$ representa la media de los valores muestrales, s representa la desviación estándar de los valores muestrales, σ representa la desviación estándar de los valores poblacionales y μ representa la media de la población.

3. Un error de tipo I es el error de rechazar la hipótesis nula cuando esa hipótesis nula en realidad es verdadera. Un error de tipo II es el error de no rechazar la hipótesis nula cuando esa hipótesis nula en realidad es falsa.
5. No tiene sentido. 7. No tiene sentido.
9. Tiene sentido. 11. No tiene sentido.
13. $z = 1.33$. La hipótesis alternativa no se respalda.
15. $z = -5.33$. La hipótesis alternativa es respaldada.
17. $z = -2.50$. La hipótesis alternativa es respaldada.
19. $z = 2.80$. La hipótesis alternativa es respaldada.
21. 0.3446. H_a no es respaldada.
23. 0.0287. H_a es respaldada.
25. 0.1096. H_a no es respaldada.
27. 0.0892. H_a no es respaldada.
29. 0.0035. H_a es respaldada.
31. 0.0179. H_a es respaldada.
33. 0.8808. H_a no es respaldada.
35. La prueba es de cola derecha, pero la puntuación estándar está a la izquierda del centro de la distribución, por lo que el valor P debe ser mayor que 0.5. La hipótesis alternativa no es respaldada. La prueba de hipótesis formal no es necesaria porque no hay manera de que podamos respaldar una afirmación de que la media es mayor que algún valor cuando la media muestral es menor que ese valor.
37. $H_0: \mu = 7.5$ años, $H_a: \mu < 7.5$ años, $n = 100$, $\bar{x} = 7.01$ años, $s = 3.74$ años, $\sigma \approx s$, $z = -1.3$, no es significativo al nivel 0.05, valor $P = 0.0968$; no hay evidencia suficiente para respaldar la afirmación de que el tiempo medio de posesión es menor a 7.5 años.
39. $H_0: \mu = 16$ onzas, $H_a: \mu \neq 16$ onzas; $n = 144$, $\bar{x} = 15.8$ onzas, $s = 1.6$ onzas, $\sigma \approx s$, $z = -1.5$, no es significativo al nivel 0.05, valor $P = 0.0668$; no hay evidencia suficiente para respaldar la afirmación de que el peso medio de los paquetes es diferente de lo anunciado de 16 onzas.
41. $H_0: \mu = 600$ mg, $H_a: \mu \neq 600$ mg; $n = 65$, $\bar{x} = 589$ mg, $s = 21$ mg, $\sigma \approx s$, $z = -4.2$, es significativo al nivel 0.05, valor $P = 0.0004$, hay evidencia suficiente para respaldar la afirmación de que la cantidad media de acetaminofén es diferente de lo listado, 600 mg.
43. $H_0: \mu = 21.4$ mpg, $H_a: \mu < 21.4$ mpg; $n = 40$, $\bar{x} = 19.8$ mpg, $s = 3.5$ mpg, $\sigma \approx s$, $z = -2.9$, es significativo al nivel 0.05, valor $P = 0.0019$; hay evidencia suficiente para respaldar la afirmación de que el millaje medio de gasolina para camionetas es menor que 21.4 mpg.
45. $H_0: \mu = 0.21$, $H_a: \mu > 0.21$, $n = 32$, $\bar{x} = 0.83$, $s = 0.24$, $\sigma \approx s$, $z = 14.6$, es significativo al nivel 0.05,

valor $P < 0.0002$; hay evidencia suficiente para respaldar la afirmación de que la media es mayor que 0.21.

47. $H_0: \mu = 3.39$ kg, $H_a: \mu \neq 3.39$ kg, $n = 121$, $\bar{x} = 3.67$ kg, $s = 0.66$ kg, $\sigma \approx s$, $z = 4.7$, es significativo al nivel 0.05, valor $P < 0.0004$; hay evidencia suficiente para respaldar la afirmación de que el peso medio al nacer de bebés hombres de mujeres que tomaron un complemento de vitaminas es diferente del promedio nacional para todos los bebés hombres.
49. $H_0: \mu = 24$ meses, $H_a: \mu < 24$ meses; $n = 70$, $\bar{x} = 22.1$ meses, $s = 8.6$ meses, $\sigma \approx s$, $z = -1.8$, es significativo al nivel 0.05, valor $P = 0.0359$; hay evidencia suficiente para respaldar la afirmación de que la media es menor que 24 meses.
51. Error de tipo I: rechazar la afirmación de que el paciente está libre de una enfermedad cuando el paciente en realidad está libre de la enfermedad. Error de tipo II: no rechazar la afirmación de que el paciente está libre de una enfermedad cuando el paciente no está libre de una enfermedad.
53. Error de tipo I: rechazar la afirmación de que la lotería es justa cuando la lotería no es justa. Error de tipo II: no rechazar la afirmación de que la lotería es justa cuando la lotería está sesgada.

SECCIÓN 9.3

1. El símbolo p representa la proporción en una población, $\hat{p}$ representa la proporción en una muestra y un valor P es la probabilidad de obtener una proporción muestral que sea al menos tan extrema como la proporción muestral que se está considerando.
3. No. La muestra es de respuesta voluntaria, no una muestra aleatoria simple. Es probable que aquellos que respondan no sean representativos de la población general.
5. Tiene sentido. 7. No tiene sentido.
9. $H_0: p = 0.5$, $H_a: p > 0.5$; $n = 400$, $\hat{p} = 0.51$, la desviación estándar de la distribución de proporciones muestrales es 0.025, $z = 0.5$, no es significativa al nivel 0.05, valor $P = 0.3446$; no existe evidencia suficiente para respaldar la afirmación de que una mayoría de votantes respalda al candidato.
11. $H_0: p = 0.50$, $H_a: p < 0.50$; $n = 1015$, $\hat{p} = 0.44$, la desviación estándar de la distribución de proporciones muestrales es 0.0157, $z = -3.8$, es significativa al nivel 0.05, valor $P < 0.0002$; existe evidencia suficiente para respaldar la afirmación de que menos de la mitad de todos los adolescentes en la población sienten que las calificaciones son la mayor fuente de presión.
13. $H_0: p = 0.099$, $H_a: p < 0.099$; $n = 1050$, $\hat{p} = 0.0933$, la desviación estándar de la distribución de proporciones muestrales es 0.0092, $z = -0.6$, es significativa al nivel

0.05, valor $P < 0.2743$; no existe evidencia suficiente para respaldar la afirmación de que el uso de drogas ilegales está por debajo del promedio nacional.

15. $H_0: p = 0.5$, $H_a: p > 0.5$; $n = 276\,000$, $\hat{p} = 0.51$, la desviación estándar de la distribución de proporciones muestrales es 0.000952, $z = 10.5$, es significativa al nivel 0.05, valor $P < 0.0002$; existe evidencia suficiente para respaldar la afirmación de que más de la mitad de todos los estudiantes de primer año consideran que el aborto debe ser legal.
17. $H_0: p = 0.53$, $H_a: p \neq 0.53$; $n = 3\,600$, $\hat{p} = 0.54$, la desviación estándar de la distribución de proporciones muestrales es 0.0083, $z = 1.2$, no es significativa al nivel 0.05, valor $P = 0.2302$; no existe evidencia suficiente para respaldar la afirmación de que ha cambiado el porcentaje de hogares que usa gas natural.

EJERCICIOS DE REPASO DEL CAPÍTULO 9

- $H_0: \mu = 12$ onzas
 - $H_a: \mu > 12$ onzas
 - $z = 10.4$
 - $z = 1.645$
 - Menos que 0.0002
 - Rechazar la hipótesis nula y la afirmación de que las latas de coca tienen una cantidad media de cola mayor que 12 onzas.
 - Un error de tipo I resultaría si concluimos que la media de la población fue mayor que 12 onzas, cuando en realidad no lo fue.
 - Un error de tipo II resultaría si no concluimos que la media población fue mayor que 12 onzas, cuando en realidad fue mayor que 12 onzas.
 - El valor P sería la probabilidad que z es mayor que 10.4 o menor a -10.4 . Como la probabilidad que z sea mayor que 10.4 es menor a 0.0002, el valor P es menor que $2 \times 0.0002 = 0.0004$.
- $H_0: p = 0.5$
 - $H_a: p > 0.5$
 - $z = 0.8$
 - $z = 1.645$
 - 0.2119
 - No existe evidencia suficiente para respaldar la afirmación de que la mayoría son fumadores un año después del tratamiento con parche de nicotina.
 - Un error de tipo I sería el error de respaldar la afirmación de que la mayoría de los sujetos están fumando un año después cuando la proporción de fumadores un año después es 0.5 (o menor)
 - Un error de tipo II sería el error de no respaldar la afirmación de que la mayoría está fumando un año después cuando la mayoría de los sujetos en realidad está fumando un año después.
 - El valor P es el doble para convertirse en 0.4238.
- La afirmación de un aumento en la verosimilitud de una niña no es respaldada. Con una proporción muestral menor que 0.5, no hay manera que se pueda respaldar la afirmación de que $p > 0.5$.
- La muestra no es aleatoria. Es posible que la familia no sea representativa de la población, por lo que los resultados están sesgados.

CUESTIONARIO DEL CAPÍTULO 9

- $p > 0.2$
- $p = 0.2$
- $p \neq 0.6$
- Cola derecha.
- $\mu = 110$
- $\mu > 110$
- No respaldar la afirmación de que la media es mayor que 110.
- Rechazar la hipótesis nula. No rechazar la hipótesis nula.
- Respaldar la afirmación de que la media es mayor que 110. No respaldar la afirmación de que la media es mayor que 10.
- Error de tipo I.

SECCIÓN 10.1

- La desviación estándar poblacional no es conocida y la población se distribuye normalmente, o la desviación estándar poblacional no es conocida y el tamaño de la muestra es mayor que 30.
- No hay diferencia. El término “distribución t ” es usado por lo común en lugar de distribución t de Student, pero los dos términos son equivalentes.
- No tiene sentido.
- Tiene sentido.
- $97.2 < \mu < 107.8$
- $14.3 \text{ pulg.} < \mu < 14.7 \text{ pulg.}$
- $\$6\,370 < \mu < \$11\,638$
 - $\$11\,638$
- $0.085 < \mu < 0.157$
 - La afirmación no parece ser válida. Ya que 0.165 gramos por milla no está incluido en el intervalo de confianza, ese valor no parece ser la media.
- $H_0: \mu = 0.3$; $H_a: \mu < 0.3$. El estadístico de prueba $t = -0.119$ no es menor que el valor de $t = -1.753$ encontrado en la tabla 10.1, por lo que no se rechaza la hipótesis nula. (Valor $P = 0.4534$). No hay evidencia suficiente para respaldar la afirmación de que la cantidad media de azúcar en todos los cereales es menor que 0.3 g.
- $H_0: \mu = 92.84$; $H_a: \mu \neq 92.84$. El estadístico de prueba $t = -0.425$ no es menor que el valor de $t = -2.093$ encontrado en la tabla 10.1, por lo que no se rechaza la hipótesis nula. (Valor $P = 0.6758$). No hay evidencia suficiente para respaldar la afirmación de que las nuevas bolas de béisbol tienen una altura media de rebote que es diferente de 92.84 pulgadas. Las nuevas bolas de béisbol no parecen ser diferentes.
- $H_0: \mu = 3.39$; $H_a: \mu \neq 3.39$. El estadístico de prueba $t = 1.735$ no es mayor que el valor de $t = 2.131$ encontrado en la tabla 10.1, por lo que no se rechaza la hipótesis nula. (Valor $P = 0.1032$). No hay evidencia suficiente para respaldar la afirmación de que el peso medio al nacer para todos los bebés hombres de madres que tomaron vitaminas es diferente de 3.39 kg. Con base en los resultados dados, el complemento vitamínico no parece tener un efecto en el peso al nacer.

- 23.** $H_0: \mu = 10.5$; $H_a: \mu < 10.5$. El estadístico de prueba $t = -0.095$ no es menor o igual al valor de $t = -1.717$ encontrado en la tabla 10.1, por lo que no se rechaza la hipótesis nula. (Valor $P = 0.4627$). No hay evidencia suficiente para respaldar la afirmación de que el tiempo medio es menor que 10.5 s. Los tiempos han sido medidos con mayor precisión en los últimos años. La prueba de hipótesis no toma en cuenta el patrón de disminución en los tiempos.

SECCIÓN 10.2

- Una tabla de dos entradas incluye las frecuencias correspondientes a dos variables diferentes. Una variable es usada para las categorías de los renglones y la otra variable es usada para las categorías de las columnas.
- No. Es común encontrar que los valores esperados no son números enteros. Son frecuencias esperadas en promedio, no frecuencias que se observarían en casos específicos.
- No tiene sentido. **7.** Tiene sentido.
- Valor crítico: $\chi^2 = 6.635$. No rechazar la hipótesis nula de independencia. Género y respuesta no parecen estar relacionados.
- Valor crítico: $\chi^2 = 3.841$. No rechazar la hipótesis nula de independencia. Género y respuesta no parecen estar relacionados.
- a.** H_0 : género y categoría del estudiante (tiempo completo o tiempo parcial) son independientes. H_a : género y categoría del estudiante (tiempo completo o tiempo parcial) están relacionados de alguna manera.
b. 48.24, 35.76, 36.76, 27.24 **c.** 0.174 **d.** 3.841
e. No rechazar la hipótesis nula de independencia. No parece haber relación entre género y categoría del estudiante (tiempo completo o tiempo parcial).
- a.** H_0 : respuesta es independiente de si el sujeto es trabajador o jefe de nivel superior. H_a : hay alguna relación entre respuesta y si el sujeto es un trabajador o un jefe de nivel superior. **b.** 181.60, 254.40, 50.40, 70.60 **c.** 4.698 **d.** 3.841 **e.** Rechazar la hipótesis nula de independencia. Parece que existe una relación entre respuesta y si el sujeto es un trabajador o un jefe de nivel superior.
- a.** H_0 : mejoría es independiente de si al participante se le dio la droga o un placebo. H_a : mejoría y tratamiento (droga o placebo) están relacionados de alguna manera.
b. 54.16, 43.84, 50.84, 41.16 **c.** 0.289 **d.** 3.841
e. No rechazar la hipótesis nula de independencia. No parece haber una relación entre mejoría y tratamiento (droga o placebo).
- a.** H_0 : el tipo de crimen es independiente de si el criminal es un desconocido. H_a : el tipo de crimen está relacionado con si el criminal es un desconocido.

- b.** 29.93, 284.64, 803.43, 21.07, 200.36, 565.57 **c.** 119.330 **d.** 9.210 **e.** Rechazar la hipótesis nula de independencia. El tipo de crimen parece que está relacionado de alguna manera con si el criminal es un desconocido.

SECCIÓN 10.3

- ANOVA representa análisis de varianza. Es un método para probar la igualdad de tres o más medias poblacionales.
 - No. Los estudiantes del Colegio de Newport no necesariamente son representativos de los estudiantes en todos los colegios de Estados Unidos.
 - No tiene sentido. **7.** No tiene sentido.
 - a.** H_0 : las páginas de los tres libros tienen la misma media de puntuación de calidad de Flesch-Kincaid.
b. H_a : las páginas de los tres libros tienen puntuaciones medias de calidad de Flesch-Kincaid que no todas son iguales. **c.** 0.00077 **d.** Rechazar la hipótesis nula de medias iguales. Las páginas de los tres libros no tienen la misma media de puntuación de calidad de Flesch-Kincaid.
 - a.** Los tres grupos de edades tienen los mismos tiempos medios. **b.** Los tres grupos de edades tienen tiempos con medias que no todas son iguales. **c.** 0.828324293
d. No. El valor P grande sugiere que no debemos rechazar la hipótesis nula de medias iguales. Las tres medias poblacionales no parecen ser diferentes.
 - Con un valor P de 0.4216, no rechazar la hipótesis nula de medias iguales. No hay evidencia suficiente para rechazar la afirmación de que las categorías de pesos diferentes tienen la misma media. Con base en estos datos muestrales, no podemos concluir que los automóviles más grandes son más seguros.
 - Con un valor P de 0.0305, rechazar la hipótesis nula de medias iguales. Existe evidencia suficiente para rechazar la afirmación de que las épocas diferentes tienen la misma media.
- ## EJERCICIOS DE REPASO DEL CAPÍTULO 10
- a.** 7.5% de aquellos que usaron un teléfono celular tuvieron un accidente, comparado con 10.2% de aquellos que no usaron un teléfono celular. Con base en estos resultados, el uso del teléfono celular no parece ser peligroso. **b.** H_0 : tener un accidente el año pasado es independiente de si la persona usa un teléfono celular. H_a : tener un accidente el año pasado y usar un teléfono celular son dependientes. **c.** 27.76, 277.24, 41.24, 411.76 **d.** 1.505 **e.** Puesto que el estadístico de prueba de 1.505 es menor que 3.841, el valor P es

mayor que 0.05. **f.** No hay evidencia suficiente para justificar el rechazo de la afirmación de que tener un accidente el año pasado es independiente de si la persona usa un teléfono celular. Con base en estos resultados, los teléfonos celulares no parecen ser un factor para tener accidentes. **g.** Más información sugiere que el uso del teléfono celular causa una distracción que se vuelve peligrosa. Algunos estados han prohibido el uso de teléfono celular a conductores.

- 2.** $H_0: \mu = 165$, $H_a: \mu > 165$. El estadístico de prueba $t = 8.398$ es mayor que el valor de $t = 1.943$ encontrado en la tabla 10.1, por lo que se rechaza la hipótesis nula. (Valor $P = 0.00008$). Hay evidencia suficiente para respaldar la afirmación de que la media de la carga axial es mayor que 165 libras.
- 3.** $227.2 \text{ lb} < \mu < 278.3 \text{ lb}$

- 4. a.** Los tres grupos de edades tienen la misma temperatura media corporal. **b.** Los tres grupos de edades tienen temperaturas medias corporales que no todas son iguales. **c.** 0.194915 **d.** No. El valor P es mayor que 0.05, lo cual sugiere que no debemos rechazar la hipótesis nula de medias iguales.

CUESTIONARIO DEL CAPÍTULO 10

- 1.** b **2.** b **3.** a **4.** $H_0: \mu = 1500$
- 5.** $H_a: \mu > 1500$ **6.** Falso. **7.** Análisis de varianza.
- 8.** No rechazar la hipótesis nula que las tres poblaciones tienen la misma media.
- 9.** H_0 : sexo y respuesta son independientes.
- 10.** Rechazar la hipótesis nula.

Índice

A

Abelson, Robert, 384
Acuerdo, método de, 316
Agencia de Protección Ambiental (EPA, por sus siglas en inglés), 39
Agencia libre, 78
Agencia para política e investigación de cuidado de la salud, 31
Agresión en niños, 48-49
Agrupación inadecuada, 301-303
Alcohol, 70, 111, 135, 262, 428
Alelos, 283
Alimentos GM, 406-408
Allen, Marty, 76
Allensworth, Stephen, 276
Altria, 36
American Statistical Association, 198
Amfotericina B, 25
Análisis de varianza, 429-432, 438
Análisis de varianza de dos factores (dos entradas), 430
ANOVA, 429-432, 438
ANOVA de un factor (o de una entrada), 429-432
A priori, método, 242
Arbuthnot, John, 373
Aristóteles, 369
Armstrong, Lance, 161
Aspirina, y cardiopatía, 35, 302-303
Astragali, 240
Automóviles
 accidentes mortales y, 111, 123, 135, 160, 260-262, 296
 bolsas de aire, 317
 contaminación y, 173
 datos de accidentes en, 434
 distancia de frenado de, 130
 millaje de, 377
Autos. *Vea* Automóviles
Avance Nacional de Valoración de la Educación (NAEP, por sus siglas en inglés), 325-327

B

Bacteria del maíz, Bt, 406
Barret, Stephen, 444
Baruch, Bernard, 421
Bayes, Thomas, 179
Benford, Frank, 440
Bernoulli, Daniel, 254
Berra, Yogi, 21, 308
Binet, Alfred, 229
BINGO, 269
Bolsas de aire, 317

Browning, Robert, 164
BTU, 91
Buffer, Warren, 189
Bulwer-Lytton, Edward, 126

C

Calentamiento global, 140-143, 328-330.
 Vea también Datos de dióxido de carbono
Calificaciones del CI, 209, 210, 229-231
Cambio absoluto, 68
Cambio relativo, 68, 69
Campo de energía humana, 443-444
Cáncer. *También Vea* Tabaquismo
 de pecho, 179-180
 de pulmón, 39, 316
 de piel, 122
 muertes por, 115
Cáncer de pecho, 179-180
Cáncer de pulmón, 39, 316
Cáncer en la piel, 122
Casa de Moneda de Estados Unidos, 206
Casas
 precios de, 80-81, 121
 propietarios de, 126
Casco en bicicleta, 427
Categorías, 90
Causalidad, 303-304, 315-319
Causa más allá de duda razonable, 318
Causa posible, 318
Causa probable, 318
Censo, 2, 11. *Vea también* Oficina de Censo, Estados Unidos
Centro de Investigación Pew para la Gente y la Prensa, 6
Centro Nacional para Estadísticas Educativas, 325, 380
Centro para Control de Enfermedades, 310
Cero absoluto, 57
Chevalier de Mire, 276
Cigarros. *Vea* Tabaquismo
Clancy, Tom, 429-432
Clase mediana, 153
Clases, 90
Clasificación, 90-93, 153, 283-284
Clon, 6
Coeficiente
 de correlación, 291, 294, 311, 324
 de determinación, 311-313
Coeficiente de correlación, 291, 294, 311-313
 uso de la tecnología para calcular, 324
Colas, 164
Cole, K. C., 374

College Board (Consejo Escolar), 226-228.
 Vea también SAT
Comisión de Estudio del Impacto Nacional del Juego, 280, 281
Compañía Gallup, 357
Compañía Pfizer, 25
Compañía Philip Morris, 36
Complemento, 244
Computadoras. *Vea también* Excel, SPSS, STATDISK
 circuitos para, 131, 162
 simulaciones con, 243, 251, 255, 330
 velocidad de, 127
 uso de, 114-115
Conclusiones, 374
 redacción de, 41, 72
Confianza, en causalidad, 318
Confusión, 25
Conteo, 418
Contornos, 116
Consumo de proteínas, 346-350
Convertidor catalítico, 173
Copa Mundial de Fútbol, 356
Correlaciones, 286-294
 interpretación, 299-304
Correlaciones negativas, 290
Correlaciones no lineales, 290
Correlaciones positivas, 290
Cotina, 110, 167-169
Cráneos egipcios, 435
Cráter producido por meteoro, 128
Criterio de Pareto, 192
Cromosomas, 282-283
Cuartil central, 166-167
Cuartiles, 166-167
Cuartil inferior, 166-167
Cuartil superior, 166-167
Curie, Marie, 195
Curva gaussiana, 200. *Vea también* Distribución normal

D

Datos, 2
 fraudulentos, 439-442
 sin procesar, 4
 tipos de, 54-55
Datos atípicos, 149-150, 299-300
Datos categóricos, 54
Datos continuos, 55
Datos cualitativos, 54
Datos cuantitativos, 54
Datos de acciones, 92, 113, 188-190, 303-304

- Datos de afiliación política, 103
 Datos de asteroides, 127-128
 Datos de básquetbol, 268
 Datos de béisbol, 78-79, 156, 296
 Datos de calificaciones, 183
 Datos de calorías, 298
 Datos de clima, 292-293, 297
 Datos de cometa, 127-128
 Datos de costo de la universidad, 128-129
 Datos de crimen, 428
 Datos de declaración de culpabilidad, 425-426
 Datos de desempleo, 4, 16, 311-312, 356
 Datos de dióxido de carbono, 101-102, 140-142, 173, 329-330
 Datos de distancia de frenado, 130
 Datos de divorcio, 121
 Datos de edad, 107-108, 131
 Datos de estándar de vida, 86-88
 Datos de estatura, 210, 307
 Datos de diamante, 286-289, 307-308, 312-313
 Datos de homicidios, 108-109
 Datos de impuestos, 84-85
 Datos de influenza, 109
 Datos de ingresos, 78-79, 82, 86-87, 121, 130, 132, 149, 191-193, 380
 Datos de internet, 114-115
 Datos de lanzamiento de moneda, 238-241, 245-246, 254-256, 268, 376
 Datos de lectura, 325-327
 Datos de límite de velocidad, 296
 Datos del Viejo Fiel, 162
 Datos de matrimonios, 121
 Datos de melanoma, 122
 Datos de migración de aves, 118-119
 Datos de nacimiento, 132, 241, 306, 315, 363, 373
 Datos de neumonía, 115
 Datos de nivel de colesterol, 210
 Datos de periódico, 123
 Datos de precios de bonos, 113
 Datos de precios de gasolina, 75-77, 78, 79
 Datos de presupuesto, 64, 122
 Datos de presupuesto federal, 64, 122
 Datos de producción de grano, 295
 Datos de riesgo de esquizofrenia, 132
 Datos de riqueza, 191-193
 Datos de ruleta, 258
 Datos de salario, 78-79, 149, 380. *Vea también* Datos de ingresos
 Datos de segregación, 123
 Datos de tamaño de granjas, 291-292, 340
 Datos de tamaño de pecho, 198
 Datos de teléfono celular, 131, 436
 Datos de televisión, 125, 204, 297, 302.
Vea también Índice de audiencia Nielsen
 Datos de tiro de dados, 215-217, 245-246, 251, 271
 Datos de tuberculosis, 115, 183
 Datos de uso de energía, 90-91, 95, 116, 130
 Datos de víctimas mortales por arma de fuego, 124
 Datos de víctimas mortales. *Vea* Automóviles, accidentes mortales y; Mortalidad, datos sobre
 Datos de viajes aéreos, 260-262, 265
 Datos de votación, 311-312
 Datos discretos, 55
 Datos geográficos, 116-117
 Datos olímpicos, 119-120, 417
 Datos sin procesar, 4
 “De”, 70-71
 DDT, 407
 Deciles, 169
 Decimales, 68
 Deflación, 77
 Demografía, 363
 De Montaigne, Michel, 26
 Departamento del Trabajo de Estados Unidos, 4, 16, 356
 Detectores de mentiras, 180-181
 Desviación estándar, 171-174
 Desviación de la media, 171
 Dewdney, A. K., 85
 Dewey, Thomas, 395
 Diaconis, Persi, 239
 Diagrama de cajas, 167, 169
 Diagrama de cajas modificado, 167
 Diagrama de cajas y bigotes, 167, 169
 Diagrama de Pareto, 101-102, 136
 Diagrama de puntos, 100, 101, 136
 Diagrama de serie de tiempo, 108-109, 136
 Diagrama de tallo, 105
 Diagrama de tallo y hojas, 105
 Diagrama de Venn, 270, 272
 Diagramas apilados, 115, 136
 Diagramas de dispersión. *Vea* Gráfica de dispersión
 Diferencia absoluta, 69, 70
 Diferencia, método de, 316
 Diferencia relativa, 69, 70
 Diferencias de género
 en áreas de universidad, 419
 en ingresos, 120-121, 192
 en inscripción a universidad, 120-121, 122, 127
 en reacciones a las drogas, 35
 en participación olímpica, 119-120
 en tasas de natalidad y mortalidad, 241
 en tiempos de carrera, 310
 en veredictos de jurado, 270
 Dígitos significativos, 64, 147
 Dirksen, Everett, 62
 Disraeli, Benjamín, 1
 Distorsiones de percepción, 125
 Distorsión, gráfica, 125-130
 Distribución
 de medias muestrales, 216-218, 334-340, 347, 380-381
 de proporciones muestrales, 340-343
 formas de, 157-161
 graficación, 99-109, 112-120
 normal. *Vea* Distribución normal
 probabilidad, 245-247
 Distribución bimodal, 158
 Distribución con sesgo negativo, 160
 Distribución con sesgo positivo, 160
 Distribución con un pico, 158
 Distribución de muestreo. *Vea* Distribución
 Distribución de probabilidad, 245-247
 Distribución en forma de campana.
 Vea Distribución normal
 Distribución normal, 159, 196-201, 204
 bivariada, 291
 vs distribución t , 410-411
 Distribución normal estándar, 204. *Vea también* Distribución normal
 Distribución normal bivariada, 291
 Distribución sesgada, 159-160
 Distribución sesgada a la derecha, 159-160
 Distribución sesgada a la izquierda, 159-160
 Distribución simétrica, 159-160
 Distribución t de Student, 410-415, 437-438
 Distribución uniforme, 157-159, 215
 Distribución unimodal, 158
 Distribución trimodal, 158
 DJIA. *Vea* Índice Industrial Dow Jones
 de distribución de las medias muestrales, 216-218, 347
 de distribución de proporciones muestrales, 343
 de la distribución normal, 198, 204
 DNA, 282-284
 Documentos Federales (Federalist Papers), 367
 Dolly, la oveja, 6
 Donaciones de aluminio, 313
 Dow, Charles H., 188
 Doyle, Arthur Conan, 93, 308
 Drogas, efectividad de, 25, 31, 35
 Duda razonable, 318
- ## E
- Educación
 y esperanza de vida, 403-405
 y tabaquismo, 428
 Efecto de la isla de calor urbana, 60, 140
 Efecto del experimentador, 27-28
 Efecto de selección, 37, 38
 Efecto Flynn, 229-231
 Efecto Hawthorne, 27
 Efecto invernadero, 328-329
 Efecto Mozart, 25
 Efecto placebo, 25-26
 Efecto Rosenthal, 28
 Efron, Bradley, 366
 Einstein, Albert, 11
 Ejercicio, y cáncer de pecho, 179-180
 Electores, 16
 Elevaciones, 117

Eliot, T. S., 262
 Encuesta Roper, 38
 Encuestas
 autoselección, 37
 redacción de, 40
 Encuestas autoseleccionadas, 37
 Encuestas de respuesta voluntaria, 37
 Enfermedad cardiovascular. *Vea*
 Enfermedad cardíaca
 Enfermedades cardíacas
 mortalidad por, 115
 y aspirinas, 35, 302-303
 y cirugía, 35, 317-318
 y salvado de avena, 304
 Enunciado de probabilidad, 211
 Error absoluto, 61-62
 Error aleatorio, 59-61
 Error de calibración, 59
 Error de disponibilidad, 40
 Error de muestreo, 338
 Error de tipo I, 388-389
 Error de tipo II, 388-389
 Errores
 absoluto *vs* relativo, 61-62
 calibración, 59
 de muestreo, 338
 disponibilidad, 40
 margen de, 5-6, 72, 347-348, 356, 357
 sistemático *vs* aleatorio, 59-61
 tipo I/II, 388-389
 Error relativo, 61-62
 Error sistemático, 59-61
 Escala exponencial, 126-127
 Escala Kelvin, 57
 Escala logarítmica, 126-127
 Escalas, 56-57, 100, 126-128
 Esperanza de vida, 70, 263-265, 290-291,
 308, 403-405
 Estadística
 descriptiva, 7
 inferencial, 7, 333
 muestral, 4-6, 187, 341
 orígenes de la, 199, 363-365
 propósito de, 8
 prueba, 382
 Estadística inferencial, 7, 333
 Estadística descriptiva, 7. *Vea también*
 Muestreo estadístico
 Estadísticas muestrales, 4-6, 341
 uso de tecnología para obtener, 187
 Estadísticas vitales. *Vea* Datos de
 natalidad; datos sobre mortalidad
 Estadístico de prueba, 382
 Estadístico ji cuadrada, 420, 422-423
 Estrato, 16
 Estudiantes. *Vea* Estudiantes universitarios
 Estudiantes universitarios
 áreas de, 103-104
 género de, 120-121, 122, 127
 salarios iniciales de, 149, 380

Estudio control de caso, 23
 Estudio de cuidado en guarderías, 48-50
 Estudio de observación, 21, 22, 23-24
 Eventos independientes, 268
 Estudio de salud de enfermeras de
 Harvard, 23, 51-52
 Estudio prospectivo, 23
 Estudio retrospectivo, 23
 Estudios estadísticos
 evaluación, 34-41
 pasos en, 7-8
 tipos de, 21-31
 Evento, 238-239
 dependiente, 268-269
 independiente, 268
 Eventos dependientes, 268-270
 Eventos mutuamente excluyentes, 270-271
 Eventos no traslapados, 270-271
 Eventos traslapados, 271-273
 Examen de Aptitud Escolar. *Vea* SAT
 Excel
 exhibición visual de datos con, 136
 exploración de correlación con, 324
 obtención de estadísticas descriptivas
 con, 187
 obtención de estimación de intervalo de
 confianza con, 362, 437
 pruebas de hipótesis con, 402, 437-438
 trabajo con puntuaciones *z* en, 225
 Experimento, 21, 22, 315
 diseño de, 24-29
 Experimento ciego (sencillo), 28
 Experimento doble ciego, 28

F

F, 431-432
 Falacia del jugador, 254-256
 Falsificación de peniques (monedas de un
 centavo), 206
 Falsos negativos, 180
 Falsos positivos, 180
 Flood, Curt, 78
 Fluconazol, 25
 Flynn, James, 229-231
 Forbes 400, 192
 Fracciones, conversión de, 68
 Franken-alimentos, 406-408
 Franklin, Benjamín, 239
 Fraude, detección, 439-442
 Frecuencia, 90, 418. *Vea también* Tablas
 de frecuencia
 acumulada, 94-95
 esperada, 421-422
 relativa, 93-95, 196, 199, 242, 420
 Frecuencia acumulada, 94-95
 Frecuencia relativa, 93-95, 196, 199, 242,
 420
 Frecuencia relativa acumulada, 94
 Frecuencias esperadas, 421-422
 Fumador activo, 167-169
 Fumador pasivo, 168-169

G

Galeno, 312
 Gallup, George, 37
 Galton, Francis, 307, 333
 Gases de invernadero, 140-142, 328-330.
 Vea también Datos de dióxido de
 carbono
 Gates, Bill, 191
 Gauss, Carl Friedrich, 200
 Gen, 282-283
 Generador de números aleatorios, 13
 Ginsburg, Ruth Bader, 318
 Gossett, William, 410
 Grados de libertad, 411-412
 Gráfica de barras, 99-102, 104-106,
 112-114, 136
 Gráfica de línea, 106-108, 113-115, 136
 Gráficas, 99-109, 112-120, 137-143
 engaño con, 125-130
 rótulos de, 100
 uso de tecnología para crear, 136
 Gráficas circulares, 103-104, 136
 Gráficas combinadas, 119
 Gráficas de dispersión, 286-293
 uso de tecnología para crear, 324
 Gráficas de múltiples barras, 112-114
 Gráficas de múltiples líneas, 113-115, 136
 Gráficas en tres dimensiones, 118
 Graham, Benjamin, 189
 Grandes números, ley de, 251-252
 Graunt Juego, 257
 Grant, John, 363-365
 Grosvenor, Charles H., 151
 Grupo de control, 24, 26
 Grupo de tratamiento, 24, 26
 Guardería,
 y abuso de niños, 28
 y niños agresivos, 48-50

H

H_0 , 371-372, 418
 H_a , 371-372, 418
 Halifax, George Savile, 235
 Halley, Edmund, 364
 Hipótesis alternativa (H_a), 371-372, 418
 Hill, Theodore P., 441
 Hipótesis, 371. *Vea también* Pruebas de
 hipótesis
 Hipótesis nula (H_0), 371-372, 418
 Histogramas, 104-106, 108, 136, 146
 Hombres. *Vea* Diferencia de género
 Hogares. *Vea* Casas
 Horóscopos, 35
 Huellas digitales, DNA, 282-284
 Huracán Floyd, 243
 Huracán Katrina, 75, 252

I

Ibuprofeno, 35
 IMC, 76
 Índice de audiencia Nielsen, 2-4, 13, 356

Índice de confianza del consumidor, 79
 Índice de costo de la risa, 79
 Índice de masa corporal (IMC), 76
 Índice de precios al consumidor (IPC), 77-79, 86-87, 131
 Índice de precios al productor, 79
 Índice Gini, 191-192
 Índice Industrial Dow Jones (DJIA), 92, 111, 188-190, 441
 Índice Nielsen de radio, 3
 Industrias ProCare, Ltd., 370
 Inflación, 77
 Ingreso de adolescentes, 112-113, 147
 Ingreso real, 86
 Ingresos. *Vea* Datos de ingresos
 Insinuación, 37
 Intervalo de confianza, 5-6, 347-351, 355-356, 411-412
 uso de tecnología para determinar, 362, 437
 Instituto de Investigación de Tabaco, 36
 Instituto Nacional de Salud Infantil y Desarrollo Humano (NICHD, por sus siglas en inglés), 48-50
 Inundación cada 500 años, 242
 IPC, 77-79, 86-87
 IPC-U, 77
 IPC-W, 77
 Irwin, Wallace, 299
 Ivins, Molly, 53

J

Jeffreys, Alec, 282
 Jones, Edward D., 188
 Jordan, Michael, 149
Journal of the American Medical Association, 12, 444
 Joyce, James, 59

K

Keynes, John Maynard, 269
 Knebel, Fletcher, 286

L

Lancet (revista), 302-303
 Landon, Alf, 37
 Ley de Benford, 440-442
 Ley de los grandes números, 251-252
 Ley de los promedios, 251-252
 Ley de Moore, 131
 Lincoln, Abraham, 334
Literary Digest (revista), 37
 Longevidad, 403-405
 Lote, 244
 Loterías, 249, 253-254, 256, 257-258, 279-281

M

Madison, James, 367
 Mamografía, 179-180
 Mapa de contornos, 116-117
 Mapas, 116-117

Marcha de la muerte de Napoleón, 137-138
 Marey, E. J., 137
 Margen de error, 5-6, 72, 347-348, 356, 357
 Mariposa monarca, 407
 “Más que”, 70-71
 Massey, William A., 219
 Media, 3, 146-153
 desviación de la, 171
 muestral, 216-218, 334-340, 380-381
 población, 334-338, 346-352, 380-390
 ponderada, 151-153
 Mediana, 146-149, 153, 167
 Media ponderada, 151-153
 Medias muestrales, 216-218, 334-340, 380-381
 Medida, niveles de, 55-57
 Metaanálisis, 30-31
 Método de los residuos, 316
 Método empírico, 242
 Método Monte Carlo, 243
 Métodos de Mill, 316
 Mill, John Stuart, 316
 Minard, Charles Joseph, 137-138
 Misión Kepler, 38
 Moda, 146-149
 Modelo a escala del Sistema Solar Voyage, 14
 Modificación genética, 406-408
 Moore, David W., 45
 Moore, Gordon E., 131
 Morfina, 35
 Mortalidad, datos sobre, 109, 241, 262-265, 290, 298, 306, 315, 363-365. *Vea también* Automóviles, accidentes mortales y en accidentes aéreos, 260-262, 265 de enfermedad, 115, 122 por armas de fuego, 124 por homicidio, 108-109
 Mortalidad infantil, 290-291, 298, 308
 Muestra, 4-5, 13-18
 Muestra aleatoria, 13-14, 16, 17
 Muestra estratificada, 16, 17
 Muestra representativa, 12
 Muestras autoseleccionadas, 15, 37
 Muestras simples aleatorias, 13-14, 16, 17
 Muestreo, 14, 16, 17
 con/sin reemplazo, 336
 Muestreo con/sin reemplazo, 336
 Muestreo de conveniencia, 15, 16, 17
 Muestreo por conglomerado, 15, 16, 17
 Muestreo sistemático, 14, 16, 17
 Mujeres. *Vea* Diferencias por género
 Munro, H. H., 61
 Museo Nacional del Aire y del Espacio, 14

N

Naciones Unidas, 37-38
 NASA, 38, 63

Nash, Ogden, 233
Natural and Political Observations on the Bills Mortality, 363
 NBC, e índice de audiencia, 13
 Newcomb, Simon, 440
New England Journal of Medicine, 35
 Newton, Isaac, 34
New York Times, 5, 39, 67, 280
 NHANES III, 346, 349
 Nicotina. *Vea* Tabaquismo
 Nielsen Media Research, 2-4, 13, 356
Nightline, 37-38
 Nigrini, Mark J., 441
 Niños
 abuso de, 28
 agresión en, 48-50
 ingreso de, 112-113
 y bolsas de aire, 317
 y tabaquismo, 168
 y televisión, 302
 Niños agresivos, en guarderías, 48-50
 Nivel de intervalo de medida, 56
 Nivel de razón de medida, 56
 Nivel de significancia. *Vea* Significancia estadística
 Nivel nominal de medida, 55-56
 Nivel ordinal de medida, 55-56
 Notación de suma, 153, 174
 Número índice, 75-79, 153
 Número variable de repeticiones en línea (VNTR, por sus siglas en inglés), 283
 Número variable de secuencias repetidas (VNTR, por sus siglas en inglés), 283

O

Oficina de Censo, Estados Unidos, 61, 64, 69, 107-108
 Orbitador climático de Marte, 63
 Oro, 113, 287, 390

P

Pago de posibilidades, 245
 Panel Intergubernamental sobre el Cambio Climático, 328
 Paradoja de Simpson, 178
 Paradojas, 178-182
 Parámetros poblacionales, 3, 4-5, 341.
 Vea también Media de la población; Proporción de la población
 Pareto, Vilfredo, 101, 192
 Participantes, 21
 Pasteur, Louis, 267
 Paxton, Charles, 366
 Peirce, Charles, 200
 Películas, éxito de, 293-294, 297
 Peligro por plomo, 310
 Peniques, 162, 206, 213
 Percentiles, 169-170, 208, 209-211
 generados por computadora, 225
 Perfil, DNA, 282-284
 Pesos de monedas, 213

Peso, y edad, 238
 Pesticidas, 406
 Peto, Richard, 302-303
 Petty, William, 363
 Pictogramas, 129-130
 Placebo, 8, 25-26
 Planetas, 38, 328-329
 Platón, 367
 Población, 3
 de Estados Unidos, 61, 64, 107-108, 156, 296
 mundial, 69, 71, 129-130, 315
 Población media, 334-338, 346-352, 380-390
 Polígono de frecuencia, 106
 Porcentajes, 67-72
 Posibilidades, 245
 Powerball (lotería), 257
 Precio índice, 75-77. *Vea también* Índice de precios al consumidor
 Precisión, 63-64
 Precisión de escanear, 244
 Premios de la academia, 96, 106
 Primer cuartil, 166-167
 Probabilidad, 238-247
 combinada, 267-273
 con números grandes, 251-256
 Probabilidad condicional, 269
 Probabilidad conjunta, 267-270
 Probabilidad de frecuencia relativa, 242, 420
 Probabilidad de strike, 243
 Probabilidades *cualquiera/o*, 270-273
 Probabilidades subjetivas, 242
 Probabilidad teórica, 239-242
 Probabilidad y (intersección), 267-270
 Problemas *al menos una vez*, 276
 Proctor, Richard, 251
 Producción de basura, 350-351, 354
 Producción de carne, 295
 Promedio, 3, 146-153. *Vea también* Media
 Promedio de bateo, 156
 Promedio de imparables, 156
 Promedio, ley de los, 251-252
 Proporción
 de la población, 341-343, 355-358, 394-398
 muestral, 341-343
 Proporción de la población, 341-343, 355-358, 394-398
 Proporciones muestrales, 340-343, 395
 Proyecto de basura, 350, 354
Proyecto de la bruja de Blair, 293
 Proyecto Phoenix (*Fénix*), 63
 Prozac, 31
 Prueba con nicorete, 278
 Prueba de cola derecha, 371, 383, 413-414
 Prueba de cola izquierda, 371, 383, 413, 414-415
 Prueba de drogas, 181-182
 Prueba de sangre de Abbot, 277

Pruebas ciegas, 28
 Pruebas de dos colas, 371, 386-387, 414
 Pruebas de hipótesis, 369-377
 para medias de la población, 380-390
 para proporciones de la población, 394-398
 con tablas de dos entradas, 418-426
 uso de tecnología para, 402, 437-438
 para la varianza, 429-432
 Pruebas del polígrafo, 180-181
 Pruebas de Raven, 230
 Pruebas de una cola, 383-385
 Prueba Stanford-Binet del CI, 209, 230
 Punto porcentual, 72
 Puntuación de facilidad de lectura de Flesch, 429-432
 Puntuaciones estándar, 208-211, 340, 381-382, 395
 generadas con computadora, 225
 tablas de, 446-448
 Puntuaciones *z*, 208-211, 340, 381-382
 de conclusiones, 41, 72
 de encuestas, 40
 generadas con computadora, 225
 tablas de, 446-448

Q
 Quetelet, Adolphe, 198
 Quilates (para diamantes), 287
 Quilates (para oro), 287
 Quintiles, 169, 192

R
 Rachas, 256
 Raciones diarias recomendadas (RDR), 346, 349
 Radón, y cáncer, 39
 Randi, James, 443-444
 Rango, 165
 Ravitch, Diane, 285
 Razón del ombligo, 98, 157
 RDR (raciones diarias recomendadas), 346, 349
 Recta de mejor ajuste, 307-312, 313
 uso de tecnología para crear, 324
 Recta de mínimos cuadrados, 313. *Vea también* Recta de mejor ajuste
 Recta de regresión, 307-312, 313
 Regla 68-95-99.7, 205-207
 Regla de Bayes, 179
 Regla de redondeo, 147
 Regla empírica para el rango, 172-173
 Regresión, 307-313
 Regresión múltiple, 312-313
 Relaciones lineales, 290
 Relaciones no lineales, 290
 “Reporte de la Nación”, 325 *Natural and Political Observations on the Bills Mortality*, 363
 Resultado, 238-239
 de una prueba de hipótesis, 373

Resumen de cinco números, 167
 Resveratrol, 30
 Revista *Discovery*, 97
 Riesgo, 260-262
 Riesgo de viaje, 260-262. *Vea también* Automóviles, accidentes mortales y
 Riesgo de VIH, 185
 Rogers, Will, 87
 Roosevelt, Franklin, 37
 Rosa, Emily, 32, 443-444
 Rose, Pete, 357
 Rowling, J. K., 429-432

S
 Sagan, Carl, 41
 Saki, 61
 Salarios, 82, 86-87, 132. *Vea también* Datos de ingresos
 Salk, Jonas, 8
 Salto de gen, 407
 Salvado de avena, 304
 SAT, 183, 205-207, 226-228, 238
 Seguridad social, 264-265
 Selección de género, 370-371, 374
 Selección de jurado, 270
 Sesgo, 12-13, 15, 35, 36-38, 389-390
 Sesgo cero, 160
 Sesgo de participación, 37
 Sesgo de publicación, 12
 Significancia práctica, 41
 Segundo cuartil, 166-167
 Sesgo de selección, 37, 38
 Shakespeare, William, 366-367, 418
 Sidransky, David, 316
 Significancia estadística, 234-236, 374-375, 383, 388-389, 396
 Simpson, Edward, 178
 Simulación, computadora, 243, 251, 255, 330
 Sinclair, John, 199
 Speizer, Frank E., 51
 Spock, Benjamín, 270
 SPSS
 exhibición visual de datos con, 136
 exploración de correlación con, 324
 obtención de estadísticas descriptivas con, 187
 obtención de estimación de intervalos de confianza con, 362, 437
 pruebas de hipótesis con, 402, 437-438
 trabajo con puntuaciones *z* en, 225
 STATDISK
 exhibición visual de datos con, 136
 exploración de correlación con, 324
 obtención de estadísticas descriptivas con, 187
 obtención de estimación de intervalo de confianza con, 362, 437
 pruebas de hipótesis con, 402, 437-438
 trabajo con puntuaciones *z* en, 225
 Suero de cotinina, 110, 167-169

Sujetos, 21
 Súper Tazón, 2-3, 303-304

T

Tabaco. *Vea* Tabaquismo
 Tabaquismo, 36, 115, 131, 167-169, 286, 316, 322, 428
 Tabla de contingencia, 418-426
 Tabla de números aleatorios, 449
 Tabla de vida, 266, 363-365
 Tablas de dos entradas, 418-426, 438
 Tablas de frecuencia, 90-95. *Vea también*
 Tablas de dos entradas
 Talmey, Paul, 342
 Tamaño de la muestra, 351-352, 357-358
 Tandy, Jessica, 106
 Tasa de accidentes, 260, 436
 Teclado Dvorak, 97
 Teclado qwerty, 97
 Telemercadeo, 37
 Temperatura
 corporal, 351, 387-388
 escalas para, 56-57
 exterior, 61, 117, 140-142, 297, 306
 Temperatura corporal, 351, 387-388
 Temperatura global promedio, 60, 140-141, 330
 Temperaturas Fahrenheit, 56, 57
 Teorema de Chebyshev, 173
 Teorema del límite central, 215-220
 Tercer cuartil, 166-167

Tercer Encuesta Nacional de Salud y
 Examen de Nutrición (NHANES III,
 por sus siglas en inglés), 346, 349
 Thisted, Ronald, 366
 Tiempos de carrera, 310, 417
 Tolstoi, León, 429-432
 Toque terapéutico (TT), 443-444
 Tratamiento, 21, 22
 Truman, Harry S., 72, 395
 TT, 443-444
 Tufte, Edward, 137
 Tukey, John, 89
 Tumores, 179-180
 Twain, Mark, 128, 263

U

Unidad térmica británica (BTU), 91

V

Vacuna contra la polio. *Vea* Vacuna de
 Salk contra la polio
 Vacuna de Salk contra la polio, 8, 22, 23, 24, 26, 28, 236
 Valor base, 75
 Valor comparado, 69, 70
 Valor crítico, 383, 412, 423, 448
 Valor de la probabilidad, *Vea* Valor *P*
 Valor de referencia, 68-69, 70, 71, 75
 Valores inusuales, 207
 Valor esperado, 252-254
 Valor nuevo, 68, 69

Valor *P*, 375, 384-385, 396, 431
 Variable de respuesta, 22, 287
 Variable explicatoria, 22, 287
 Variables, 22
 de confusión, 25, 39
 explicatoria, 22, 287
 de interés, 22, 38-39
 de respuesta, 22, 287
 Variables de confusión, 25, 39
 Variación, 160-161, 164-174
 Variación concomitante, método de, 316
 Varianza, 171, 174
 análisis de varianza, 429-432, 438
 Verdaderos negativos, 180
 Verdaderos positivos, 180
 Vitamina C, 425

W

Wall Street Journal, 92, 188-189
 Weiss, Rick, 303
 Wells, H. G., 2
 Western Electric, 27
 Wilmut, Ian, 6
 Wilson, Flip, 261
 Woolf, Virginia, 145

Y

Yankelovich Monitor, 358
Yankelovich Monitor (compañía), 358
 Yule, George, 178



Razonamiento estadístico enseña a los estudiantes como hacerse llegar más y mejor información, mostrando el papel de la estadística en muchos aspectos de la vida cotidiana.

Este texto está basado en ejemplos reales y estudios de casos para construir una comprensión de las ideas fundamentales de la estadística que se pueden aplicar a una variedad de escenarios de la vida diaria.

Los autores incluyen datos de fuentes reales para ayudar a los estudiantes a ser responsables y mejores pensadores críticos, si deciden iniciar un nuevo negocio, planificar su futuro financiero, o simplemente mantenerse informados del acontecer diario.

PEARSON

Visítenos en:
www.pearsoneducacion.net

ISBN 978-607-32-0759-1

